BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

-- BRIEF DESCRIPTION OF THE DRAWINGS

This invention will be more fully described with reference to the drawings in which:

Figure 1 depicts the universal code equivalent of the mitochondrial I-*Sce*I gene (SEQ ID NOS:3 AND 4).

Figure 2 depicts the nucleotide sequence of the invention encoding the enzyme I-*Sce*I and the amino acid sequence of the natural I-*Sce*I enzyme (SEQ ID NOS: 5 and 2).

Figure 3 depicts the I-*Sce*I recognition sequence and indicates possible base mutations in the recognition site and the effect of such mutations on stringency of recognition (SEQ ID NOS: 6, 7, AND 8).

Figure 4A-B is the nucleotide sequence and deduced amino acid sequence of a region of plasmid pSCM525. The nucleotide sequence of the invention encoding the enzyme I-*Sce*I is enclosed in the box (SEQ ID NOS: 9 through 16).

Figure 5 depicts variations around the amino acid sequence of the enzyme I-*Sce*I (SEQ. ID NO:2).

Figure 6 shows Group I intron encoding endonucleases and related endonucleases (SEQ ID NOS: 17 through 44).

Figure 7 depicts yeast expression vectors containing the synthetic gene for I-*Sce*I.

Figure 8 depicts the mammalian expression vector PRSV I-*Sce*I.

Figure 9 is a restriction map of the plasmid pAF100. (See also YEAST, 6:521-534, 1990, which is relied upon and incorporated by reference herein).

Figures 10A and 10B show the nucleotide sequence and restriction sites of regions of the plasmid pAF100 (SEQ ID NOS:45 through 50). --

Figure 11 depicts an insertion vector pTSMω, pTKMω, and pTTω containing the I-*Sce*I site for *E. coli* and other bacteria.

Figure 12 depicts an insertion vector pTYW6 containing the I-*Sce*I site for yeast.

Figure 13A-C depicts an insertion vector PMLV LTR SAPLZ containing the I-*Sce*I site for mammalian cells.

Figure 14A-B depicts a set of seven transgenic yeast strains cleaved by I-*Sce*I. Chromosomes from FY1679 (control) and from seven transgenic yeast strains with I-*Sce*I sites inserted at various positions along chromosome XI were treated with I-*Sce*I. DNA was electrophoresed on 1% agarose (SeaKem) gel in 0.25 X TBE buffer at 130 V and 12°C on a Rotaphor apparatus (Biometra) for 70 hrs using 100 sec to 40 sec decreasing pulse times. (A) DNA was stained with ethidium bromide (0.2 µg/ml) and transferred to a Hybond N (Amersham) membrane for hybridization. (B) ^{32}P labelled cosmid pUKG040 which hybridizes with the shortest fragment of the set was used as a probe. Positions of chromosome XI and shorter chromosomes are indicated.

Figure 15A-E depicts the rationale of the nested chromosomal fragmentation strategy for genetic mapping. (A) Positions of *I-Sce*I sites are placed on the map, irrespective of the left/right orientation (shorter fragments are arbitrarily placed on the left). Fragment sizes as measured from PFGE (Fig. 14A) are indicated in kb (note that the sum of the two fragment sizes varies slightly due to the limit of precision of each measurement). (B) Hybridization with the probe that hybridizes the shortest fragment of the set determines the orientation of each fragment (see Fig. 14B). Fragments that hybridize with the probe (full lines) have been placed arbitrarily to the left. (C) Transgenic yeast strains have been ordered with increasing sizes of hybridizing chromosome fragments. (D) Deduced *I-Sce*I map with minimal and maximal size of intervals indicated in kb (variations in some intervals are due to limitations of PFGE measurements). (E) Chromosome subfragments are used as probes to assign each cosmid clone to a given map interval or across a given *I-Sce*I site.

Figure 16A-C depicts mapping of the *I-Sce*I sites of transgenic yeast strains by hybridization with left end and right end probes of chromosome XI. Chromosomes from FY1679 (control) and the seven transgenic yeast strains were digested with *I-Sce*I. Transgenic strains were placed in order as explained in Fig. 15. Electrophoresis conditions were as in Fig. 14. ^{32}P labelled cosmids pUKG040 and pUKG066 were used as left end and right end probes, respectively.

Figure 17A-H depicts mapping of a cosmid collection using the nested chromosomal fragments as probes. Cosmid DNAs were digested with EcoRI and electrophoresed on 0.9% agarose (SeaKem) gel at 1.5 V/cm for 14 hrs, stained with ethidium bromide and transferred to a Hybond N membrane. Cosmids were placed in order from previous hybridizations to help visualize the strategy. Hybridizations were carried out serially on three identical membranes using left end nested chromosome fragments purified on PFGE (see Fig. 16) as probes. A: ethidium bromide staining (ladder is the BRL "1kb ladder"), B: membrane #1, probe: Left tel to A302 site,

C: membrane #1, probe: Left tel to M57 site, D: membrane #2, probe: Left tel to H81 site, E: membrane #2, probe: Left tel to T62 site, F: membrane #3, probe: Left tel to G41 site, G: membrane #3, probe: Left tel to D304 site, H: membrane #3, probe: entire chromosome XI.

Figure 18 depicts a map of the yeast chromosome XI as determined from the nested chromosomal fragmentation strategy. The chromosome is divided into eight intervals (with sizes indicated in kb, see Fig. 15D) separated by seven *I-Sce*I sites (E40, A302 ...). Cosmid clones falling either within intervals or across a given *I-Sce*I site are listed below intervals or below interval boundaries, respectively. Cosmid clones that hybridize with selected genes used as probes are indicated by letters (a-i). They localize the gene with respect to the *I-Sce*I map and allow comparison with the genetic map (top).

Figure 19A-B depicts diagrams of successful site directed homologous recombination experiments performed in yeast.

Figure 20A-E. Experimental design for the detection of ~~HR~~ homologous recombination (HR) induced by I-*Sce* I. a) Maps of the 7.5 kb *tk -PhleoLacZ* retrovirus (G-MtkPL) and of the 6.0 kb *PhleoLacZ* retrovirus (G-MPL), SA is splice acceptor site. G-MtkPL sequences (from G-MtkPL virus) contains *PhleoLacZ* fusion gene for positive selection of infected cells (in phleomycin-containing medium) and *tk* gene for negative selection (in gancyclovir-containing medium). G-MPL sequences (from G-MPL virus) contains only *PhleoLacZ* sequences. b) Maps of proviral structures following retroviral integration of G-MtkPL and G-MPL. I-*Sce* I *PhleoLacZ* LTR duplicates, placing I-*Sce* I *PhleoLacZ* sequences in the 5' LTR. The virus vector (which functions as a promoter trap) is transcribed (arrow) by a flanking cellular promoter, P. c) I-*Sce* I creates two double strand breaks (DSBs) in host DNA liberating the central segment and leaving broken chromosome ends that can pair with the donor plasmid, pVRneo (d).
e) Expected recombinant locus following HR.

Figure 21A-B. A. Scheme of pG-MPL. SD and SA are splice donor and splice acceptor sites. The structure of the unspliced 5.8 kb (genomic) and spliced 4;2 kb transcripts is shown below. Heavy bar is ^{32}P radiolabelled *LacZ* probe (**P**). B. RNA Northern blot analysis of a pG MLP transformed ψ-2 producer clone using polyadenylated RNA. Note that the genomic and the spliced mRNA are produced at the same high level.

Figure 22A-B. A. Introduction of duplicated I-*Sce* I recognition sites into the genome of mammalian cells by retrovirus integration. Scheme of G-MPL and G-MtkPL proviruses which illustrates positions of the two LTRs and pertinent restriction sites. The size of *Bcl* I fragments and of I-*Sce* I fragments are indicated. Heavy bar is ^{32}P radiolabelled *LacZ* probe (**P**). **B**. Southern blot analysis of cellular DNA from NIH3T3

fibroblasts cells infected by G-MtkPL and PCC7-S multipotent cells infected by G-MPL. *Bcl* I digests demonstrating LTR mediated *PhleoLacZ* duplication; I-*Sce* I digests demonstrating faithful duplication of I-*Sce* I sites.

Figure 23A-B. Verification of recombination by Southern. A.: Expected fragment sizes in kilobase pairs (kb) of provirus at the recombinant locus. 1) the parental proviral locus. Heavy bar (**P**) is ^{32}P radioactively labelled probe used for hybridization. 2) a recombinant derived after cleavage at the two I-*Sce* I sites followed by gap repair using pVR neo (double-site homologous recombination, DsHR). 3) a recombination event initiated by the cleavage at the I-*Sce* I sites in the left LTR (single-site homologous recombination, SsHR). B.: Southern analysis of DNA from NIH3T3/G-MtkPL clones 1 and 2, PCC7-S/G-MPL clones 3 and 4 and transformants derived from cotransfection with pCMV(I-*Sce* I+) and pVRneo (1a, 1b, 2a, 3a, 3b and 4a). *Kpn* I digestion of the parental DNA generates a 4.2 kb fragment containing *LacZ* fragment. Recombinants 1a and 3a are examples of DsHR Recombinants 1b, 2a, 3b and 4a are examples of SsHR.

Figure 24A-B. Verification of recombination by Northern blot analyses. A.: Expected structure and sizes (in kb) of RNA from PCC7-S/G-MPL clone 3 cells before (top) and after (bottom) I-*Sce* I induced HR with pVRneo.I Heavy bars **P1** and **P2** are ^{32}P radioactively labelled probes. B.: Northern blot analysis of the PCC7-S/G-MPL clone 3 recombinant (total RNA). Lane 3 is parental cells, lane 3a recombinant cells. Two first lanes were probed with *LacZ* P1, two last lanes are probed with *neo* P2. parental PCC7-S/G-MPL clone 3 cells express a 7.0 kb *LacZ* RNA as expected of trapping of a cellular promoter leading to expression of a cellular-viral fusion RNA. The recombinant clone does not express this *Lacz* RNA but expresses a *neo* RNA of 5.0 kb, corresponding to the size expected for an accurate replacement of *PhleoLacZ* by *neo* gene.

Figure 25A-C. Types of recombination events induced by I-*Sce* I DSBs, a) Schematic drawing of the structure of the recombination substrate. The G-MtkPL has provirus two LTRs, each containing an I-*Sce* I recognition site and a *PhleoLacZ* gene. The LTRs are separated by viral sequences containing the *tk* gene. The phenotype of G-MtkPL containing cells is $Phleo^R$, Gls^s, β-Gal±. b) Possible modes of intra-chromosomal recombination. 1) The I-*Sce* I endonuclease cuts the I-*Sce* I site in the 5'LTR. The 5' part of U3 of the 5'LTR can pair and recombine with it homologous sequence in the 3'LTR (by single-strand annealing (SSA). 2) The I-*Sce* I endonuclease cuts the I-*Sce* I site in the 3'LTR. The 3' part of U3 of the 3'LTR can pair and recombine with its homologous sequence in the 5'LTR (by SSA). 3) The I-*Sce* I endonuclease cuts I-*Sce* I sites in the two LTRs. The two free ends can relegate (by an end-joining mechanism). The resulting recombination product in each of the three models is a solitary LTR (see right side). No modification would occur in the cellular sequences flanking the integration site. c) The I-*Sce* I endonuclease cuts the I-*Sce* I sites in the

two LTRs. The two free ends can be repaired (by a gap repair mechanism) using the homologous chromosome. On the right, the resulting recombination product is the deletion of the proviral integration locus.

Figure 26A-C. Southern blot analysis of DNA from NIH3T3/G-MtkPL 1 and 2, and *PhleoLacZ*$^-$ recombinants derived from transfections with pCMV(I-*Sce* I+) selected in Gancyclovir containing medium. a) Expected fragment sizes in kilobase pair (kbp) of parental provirus after digestion with *Pst* I endonuclease. *Pst* I digestion of the parental DNA NH3T3/G-MtkPL 1 generates two fragments of 10 kbp and of the parental NIH3T3/G-MtkPL 2 two fragments of 7 kbp and 9 kbp. b) Southern blot analysis of DNA digested by *Pst* I from NIH3T3/G-MtkPL 1, and recombinants derived from transfection with pCMV(I-*Sce* I+) (1.1 to 1.5). c) Southern blot analysis of DNA digested by *Pst* I from NIH3T3/G-MtkPL 2, and recombinants derived from transfection with pCMV(I-*Sce* I+) (2.1 to 2.6). Heavy bar is ^{32}P radiolabelled *Lac*Z probe (P).

Figure 27A-C. Southern blot analysis of DNA from NIH3T3/G-MtkPL 1 and 2, and *PhleoLacZ*$^+$ recombinants derived from transfections with pCMV(I-*Sce* I+) and pCMV(I-*Sce* I-) and selection in Phleomycin and Gancyclovir containing medium. a1) Expected fragment sizes in kbp of parental provirus after digestion with *Pst* I or *Bcl* I endonuclease. *Pst* I digestion of the parental DNA NIH3T3/G-MtkPL 1 generates two fragments of 10 kbp. *Bcl* I digestion of the parental DNA NIH3T3/G-MtkPL 2 generates three fragments of 9.2 kbp, 7.2 kbp and 6.0 kbp. a2) Expected fragment sizes in kbp of recombinants after digestion with *Pst* I or *Bcl* I endonuclease. *Pst* I digestion of DNA of the recombinant derived from NIH3T3/G-MtkPL 1 generates one fragment of 13.6 kbp. *Bcl* I digestion of the DNA of the recombinants derived from NIH3T3/G-MtkPL 2 generates two fragments of 9.2 kbp and 6.0 kbp. b) Southern blot analysis of DNA from NIH3T3/G-MtkPL 1, and recombinants derived from transfection with pCMV(I-*Sce* I-) and pCMV(I-*Sce* I+) (1c, 1d). c) Southern analysis of DNA from NIH3T3/G-MtkPL 2, and transformants derived from transfection with pCMV(I-*Sce* I-) (2a, 2b) and pCMV(I-*Sce* I+) (2c to 2h). Heavy bar is ^{32}P radiolabelled *Lac*Z probe (P).

Figure 28. Figure 28 is a diagram illustrating the loss of heterozygosity by the insertion or presence of an I-*Sce* I site, expression of the enzyme I-*Sce* I, cleavage at the site, and repair of the double strand break at the site with the corresponding chromatid.

Figure 29. Figure 29 is a diagram illustrating conditional activation of a gene. An I-*Sce* I site is integrated between tandem repeats, and the enzyme I-*Sce* I is expressed. The enzyme cleaves the double stranded DNA at the I-*Sce* I site. The double strand break is repaired by single stand annealing, yielding an active gene.

Figure 30. Figure 30 is a diagram illustrating one step rearrangement of a gene by integration of an I-*Sce* I site or by use of an I-*Sce* I site present in the gene. A

plasmid having either one I-*Sce* I site within an inactive gene, or two I-*Sce* I sites at either end of an active gene without a promoter, is introduced into the cell. The cell contains an inactive form of the corresponding gene. The enzyme I-*Sce* I cuts the plasmid at the I-*Sce* I sites, and recombination between the chromosome and the plasmid yields an active gene replacing the inactive gene.

Figure 31. Figure 31 is a diagram illustrating the duplication of a locus. An I-*Sce* I site and a distal part of the locus are inserted into the gene by classical gene replacement. The I-*Sce* I site is cleaved by I-*Sce* I enzyme, and the break is repaired by homologous sequences. This results in duplication of the entire locus.

Figure 32. Figure ~~30~~ 32 is a diagram illustrating the deletion of a locus. Two I-*Sce* I sites are added to flank the locus to be deleted. The I-*Sce* I enzyme is expressed, and the sites are cleaved. The two remaining ends recombine, deleting the locus between the two I-*Sce* I sites.

Figure 33. Figure 33 is a diagram of plasmid pG-MtkΔPAPL showing the restriction sites. The plasmid is constructed by deletion of the polyadenylation region of the tk gene from the pGMtkPL plasmid. --

On page 18, replace the last paragraph with the following amended paragraph:

-- The enzyme I-*Sce*I has a known recognition site. (ref. 14.) The recognition site of I-*Sce*I is a non-symmetrical sequence that extends over 18 bp as determined by systematic mutational analysis. The sequence reads: (arrows indicate cuts)

```
                ↓
5' TAGGGATAA CAGGGTAAT 3'   (SEQ ID NO:51)
3' ATCCC TATTGTCCCATTA 5'   (SEQ ID NO:52)
     ↑                                   --
```

On pages 47-48, replace the paragraph bridging these pages with the following amended paragraph:

-- -e- The supernatant of this clone was used to infect other mouse cells (1009) by spreading 10^5 virus particles on 10^5 cells in DMEM medium with 10% fetal calf serum and 5 mg/ml of ~~"polybrain"~~ polybrene (hexadimethrine bromide). Medium was replaced 6 hours after infection by the same fresh medium. --

On pages 71-72, replace the paragraph bridging these pages with the following amended paragraph:

-- This example describes the use of the I-*Sce* I meganuclease (involved in intron homing of mitochondria of the yeast *Saccharomyces cerevisiae*) (6B, 28B) to induce DSB and mediate recombination in mammalian cells. I-*Sce* I is a very rare-cutting restriction endonuclease, with an 18 bp long recognition site (29B, 22B). *In vivo*, I-*Sce* I endonuclease can induce recombination in a modified yeast nucleus by initiating a specific DBS leading to gap repair by the cell (30B, 17B, 21B). Therefore, this approach can potentially be used as a means of introducing specific DSB in chromosomal target DNA with a view to manipulate chromosomes in living cells. The I-*Sce* I-mediated recombination is superior to recombinase system [11B] for chromosome engineering since the latter requires the presence of target sites on both host and donor DNA molecules, leading to reaction that is reversible. --

On pages 72-73, replace the paragraph bridging these pages with the following amended paragraph:

-- pG-MtkPL was obtained in five steps: (I) insertion of the 0.3 kbp *Bgl II-Sma* I fragment (treated with Klenow enzyme) of the Moloney Murine Leukemia Virus (MoMuLV) *env* gene (25B) containing a splice acceptor (SA) between the *Nhe* I and *Xba* I sites (treated with Klenow enzyme), in the U3 sequence of the 3'LTR of MoMuLV, in an intermediate plasmid. (II) Insertion in this modified LTR of a 3.5 kbp *Nco* I-*Xho* I fragment containing the *PhleoLacZ* fusion gene [13B] (from pUT65; Cayla Laboratory, Zone Commerciale du Gros, Toulouse, France) at the *Xba* I site next to SA. (III) Insertion of this 3'LTR (containing SA and *PhleoLacZ*), recovered by *Sal* I-*EcoR* I double digestion in the p5'LTR plasmid (a plasmid containing the 5'LTR up to the nucleotide no 563 of MoMuLV [12B]) between the *Xho* I and the *EcoR* I site. (IV) Insertion of a *synthetic* I-*Sce* I recognition site into the *Nco* I site in the 3'LTR (between SA and *PhleoLacZ*), and (V) insertion (antisense to the retroviral genome) of the 1.6 kbp *tk* gene with its promoter with linker adaptators at the *Pst* I site of pG-MPL. --

On page 80, replace the paragraph beginning on line 5 with the following amended paragraph:

-- The generation of $tk^-/PhleoLacZ^-$ cells is probably a consequence of either a homo-allelic and/or an ectopic gene conversion event (36B). Isolation and detailed molecular analysis of the proviral integration sites will provide information on the relative frequency of each of these events for the resolution of chromosomal DSBs by the cell. This quantitative information is important as, in mammalian cells, the high redundancy

of genomic sequences raises the possibility of a repair of DSBs by ectopic homologous sequences. Ectopic recombination for repair of DSBs may be involved in genome shaping and diversity in evolution [29]. --

After making the above amendments, please move pages 50-58 and insert these pages after page 88 and before page 89 of the application.

Renumber pages 50-58 as pages 80-88, respectively.

Renumber pages 59-88 as pages 50-79, respectively.

On pages 89-93, replace the section entitled "References" with the following amended section.

-- References

1 B. Bernstein, N., Pennell, N., Ottaway, C. A. and Shulman, M.J. 1992. Gene replacement with one-sided homologous recombination. Mol. Cell Biol. 12: 360-367.

2 B. Bonnerot, C., Legouy, E., Choulika, A. and Nicolas, J.-F. 1992. Capture of a cellular transcriptional unit by retrovirus: mode of provirus activation in embryonal carcinoma cells. J. Virol. 66: 4982-4991.

3 B. Bonnerot, C., and Nicolas, J.-F. 1993. Application of LacZ gene fusions to post-implantation development. In "Methods in Enzymology: Guide to techniques in mouse development". 451-469. Wassarman, P.M., DePamphilis, M.L.

4 B. Brenner, D.A., Smogocki, A. and Camerini-Otero, R.D. 1986. Double-strand gap repair results in homologous recombination in mouse L cells. Proc. Natl. Acad. Sci. USA. 83: 1762-1766.

5 B. Capecchi, M.R. 1989. Altering the genome by homologous recombination. Science. 244: 1288-1292.

6 B. Jacquier, A. and Dujon, B. 1985. An intron encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell. 41: 383-394.

7 B. Jakob, H. and Nicolas, J.F. 1987. Mouse Tertocarcinoma Cells. In "Methods in Enzymology: Cell lines for genetic analysis". 66-81. Gottesman, M.H., ed., Academic Press.

8 B. Jessberger, R. and Berg, P. 1991. Repair of deletions and double-strand gaps by homologous recombination in a mammalian in vitro system. Mol. Cell Biol. 11: 445-457.

9 B. Kilby, N.J., Snaith, M.R. and Murray, J.A.H. 1993. Site-specific recombinases: tools for genome engineering. Reviews. 9: 413-421.

10 B. Lin, F.L.M., Sperle, K. and Sternberg N. 1990. Repair of double-stranded DNA breaks by homologous DNA fragments during transfer of DNA into mouse L cells. Mol. Cell Biol. 10:113-119.

11 B. Lin, F.L.M., Sperle, K. and Sternberg N. 1990. Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. Mol. Cell Biol. 10: 103-112.

12 B. Lin, F.L.M., Sperle, K. and Sternberg N. 1990. Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. Mol. Cell. Biol. 10: 103-112.

13 B. Mann, R., Mulligan, R.C. and Baltimore, D. 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell. 33: 153-160.

14 B. Mansour, S.L., Thomas K.R. and Capecchi, M.R. 1988. Disruption of the proto-oncogène int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable gene. Nature. 336:348-352.

15 B. Mulsant, P., Gatignol, A., Dolens, M. and Tiraby, G. 1988. Somat. Cell. Mol. Genet. 14: 243-252.

16 B. Nicolas, J. F. and Rubenstein, J. 1987. Retroviral vectors. Boston London Durban Singapore Sydney Toronto Wellington, Butterworths.

17 B. Plessis, A., Perrin, A., Haber, J.E. and Dujon, B. 1992. Site specific recombination determined by I-*Sce* I, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. Genetics 130:451-460.

18 B. Sauer, B. and Henderson, N. 1988. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Prac. Natl. Acad. Sci. USA. 85: 5166-5170.

19 B. Seidman, M.M. 1987. Intermolecular homologous recombination between transfected sequences in mammalian cells is primarily nonconservative. Mol. Cell. Biol. 7: 3561-3565.

20 B. Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. and Kucherlapati, R.S. 1985. Insertion of DNA sequences into the human chromosomal B-globin locus by homologous recombination. Nature. 317: 230-234.

21 B. Szostak, J. W., Orr-Weaver, T.L. and Rothstein, R.J. 1983. The double-strand break repair model for recombination. Cell. 33: 25-35.

22 B. Thierry, A., Perrin, A., Boyer, J., Fairhead, C., Dujon, B., Frey, B. and Schmitz, G. 1991. Cleavage of yeast and bacteriophage T7 genomes at a single site using the rare cutter endonuclease I-*Sce* I. Nucleic Acids Res. 19: 189-90.

23 B. Tybulewicz, V.L. J., Crawford, C.E., Jackson, P.K., Bronson, R.T. and Mulligan, R.C. 1991. Neonatal Lethality and Lymphopenia in Mice with a Homozygous Disruption of the c-abl Proto-Oncogene. Cell 65: 1153-1163.

24 B. Varmus, H. and Brown, P. 1989. Retroviruses in Mobile DNA, 58-108 (Douglas E. Berg and Martha H. Howe eds).

25 B. Weiss, R., Teich, N., Varmus, H. and Coffin, J. 1985. RNA tumor viruses. Molecular Biology of tumor viruses. Second Edition. 2) Supplements and appendixes. Cold Spring Harbor Laboratory. 1-1222.

26 B. Weiss, R., Teich, N., Varmus, H. and Coffin, J. 1985. RNA tumor viruses. Molecular Biology of tumor viruses. Second Edition. 2) Supplements and appendixes. Cold Spring Harbor Laboratory. 1-1222.

27 B. Phillips J. and Morgan W. 1994. Illegitimate recombination induced by DNA double-strand breaks in mammalian chromosomes. Molecular and Cellular Biology 14:5794-5803.

28 B. Dujon B. 1989. Group I introns are mobile genetic elements: facts and mechanistic speculations-a review. Gene 82:91-114.

29 B. Colleaux L., D'Aurio L., Galibert F. and Dujon B. 1988. Recognition and cleavage site of the intron-encoded omega transposase. Proc Natl Acad Sci USA 85:6022-6.

30 B. Fairchild C. and Dujon B. Consequences of unique double-stranded breaks in yeast chromosomes: death or homozygosis. Molecular general genetics 240:170-180.

31 B. Pfeiffer P., Thode S., Hancke J. and Vielmetter W. 1994. Mechanism of overlap information in nonhomologous DNA end joining. Molecular and Cellular Biology 14:888-895.

32 B. Mezard C. and Nicholas A. 1994. Homologous, homeologous, and illegitimate repair of double-strand breaks during transformation of a wild-type strain and a rad52 Mutant strain of Saccharomyces cerevisiae. Molecular and Cellular Biology 14:1278-1292.

33 B. Feaver W.J., Svejstrup J.Q., Bradwell L., Bradwell A.J., Buratowski S., Gulyas K., Donahue T.F., Friedberg E.C. and Kornberg R.D. 1993. Dual Roles of a Multiprotein Complex from S. cerevisiae in transcription and DNA Repair. Cell 75:1379-1387.

34 B. Kramer K., Brock J., Bloom K., Moore K. and Haber J. 1994. Two different types of double-strand breaks in Saccharomyces ceerevisiae are repaired by similar RAD52 independent, nonhomolgous recombination events. Molecular and Cellular Biology 14:1293-1301.

35 B. Weiffenbach B. and Haber J. 1981. Homothallelic mating type switching generates lethal chromosomes breaks in rad52 strains of Saccharomyces cerevisiae. Molecular and Cellular Biology 1:522-534.

36 B. Nassif N., Penney J., Pal S., Engels W. and Gloor G. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Molecular and cellular biology 14:1643-1625.

37 B. Charlesworth B., Sniegowski P. and Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eucaryotes. Nature 371:215-220. --

After page 93, and before page 94, insert the attached pages titled "SEQUENCE LISTING".

FIGURE 1

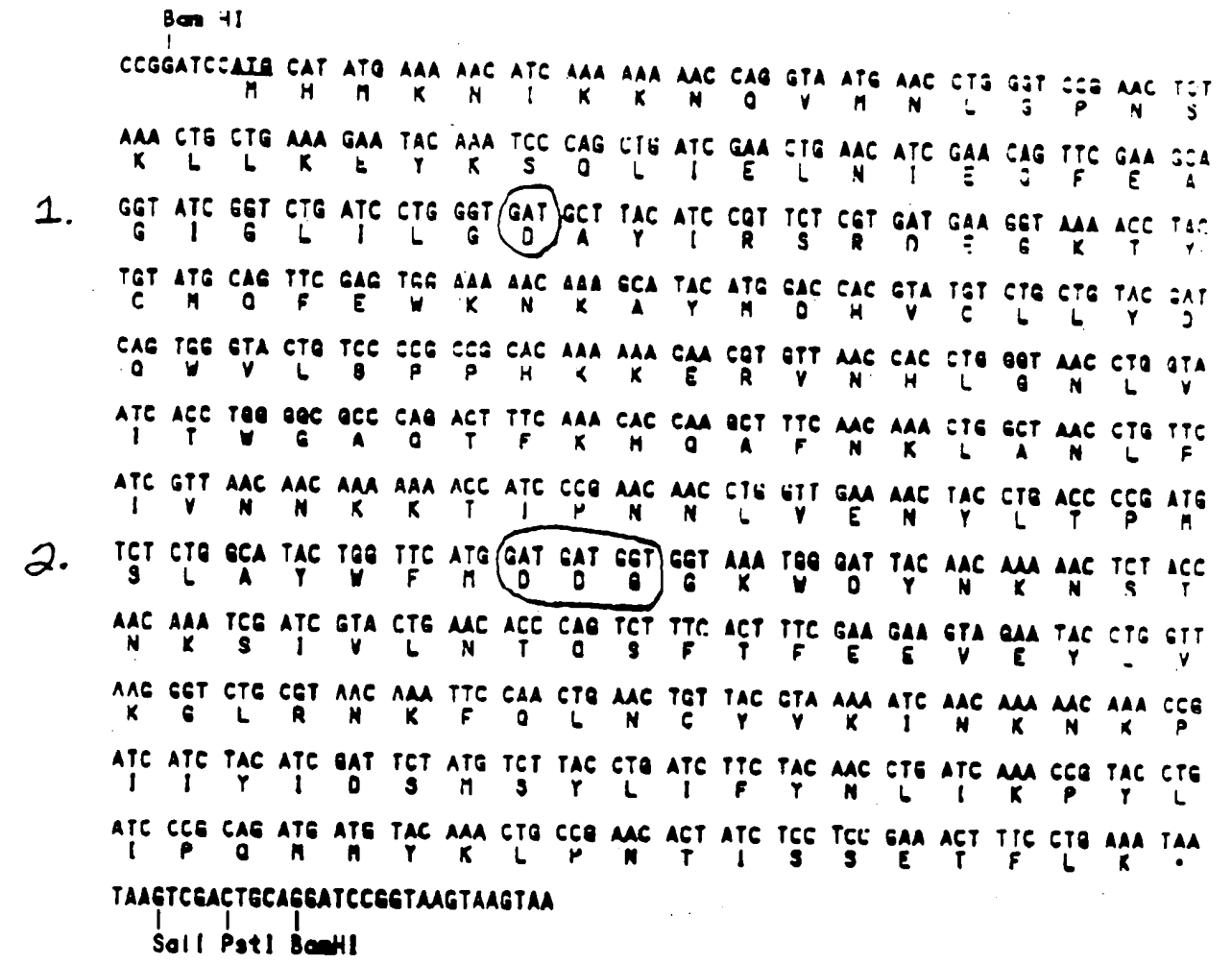
The Universal Code Equivalent of

the Mitochondrial I-*Sce* I Gene.

```
AAAAATAAAATCAT ATG AAA AAT ATT AAA AAA AAT CAA GTA ATG AAT CTC CGT CCT AAT TCT
               M   K   N   I   K   K   N   Q   V   M   N   L   R   P   N   S

AAA TTA TTA AAA GAA TAT AAA TCA CAA TTA ATT GAA TTA AAT ATT GAA CAA TTT GAA GCA
K   L   L   K   E   Y   K   S   Q   L   I   E   L   N   I   E   Q   F   E   A

GGT ATT GGT TTA ATT TTA GGA GAT GCT TAT ATT CGT AGT CGT GAT GAA GGT AAA ACT TAT
G   I   G   L   I   L   G   D   A   Y   I   R   S   R   D   E   G   K   T   Y

TGT ATG CAA TTT CAC TGC AAA AAT AAG GCA TAC ATG GAT CAT GTA TGT TTA TTA TAT GAT
C   M   Q   F   E   W   K   N   K   A   Y   M   D   H   V   C   L   L   Y   D

CAA TGG GTA TTA TCA CCT CCT CAT AAA AAA GAA AGA GTT AAT CAT TTA GGT AAT TTA GTA
Q   W   V   L   S   P   P   H   K   K   E   R   V   N   H   L   G   N   L   V

ATT ACC TGG GGA GCT CAA ACT TTT AAA CAT CAA GCT TTT AAT AAA TTA GCT AAC TTA TTT
I   T   W   G   A   Q   T   F   K   H   Q   A   F   N   K   L   A   N   L   F

ATT GTA AAT AAT AAA AAA CTT ATT CCT AAT AAT TTA GTT GAA AAT TAT TTA ACA CCT ATG
I   V   N   N   K   K   L   I   P   N   N   L   V   E   N   Y   L   T   P   M

AGT CTG GCA TAT TGG TTT ATG GAT GAT GGA GGT AAA TGG GAT TAT AAT AAA AAT TCT CTT
S   L   A   Y   W   F   M   D   D   G   G   K   W   D   Y   N   K   N   S   L

AAT AAA AGT ATT GTA TTA AAT ACA CAA AGT TTT ACT TTT GAA GAA GTA GAA TAT TTA CTT
N   K   S   I   V   L   N   T   Q   S   F   T   F   E   E   V   E   Y   L   V

AAA GGT TTA AGA AAT AAA TTT CAA TTA AAT TGT TAT GTT AAA ATT AAT AAA AAT AAA CCA
K   G   L   R   N   K   F   Q   L   N   C   Y   V   K   I   N   K   N   K   P

ATT ATT TAT ATT GAT TCT ATG AGT TAT CTG ATT TTT TAT AAT TTA ATT AAA CCT TAT TTA
I   I   Y   I   D   S   M   S   Y   L   I   F   Y   N   L   I   K   P   Y   L

ATT CCT CAA ATG ATG TAT AAA CTG CCT AAT ACT ATT TCA TCC GAA ACT TTT TTA AAA TAA
I   P   Q   M   M   Y   K   L   P   N   T   I   S   S   E   T   F   L   K   .
```

FIGURE 2

The synthetic I-Sce I gene



1 and 2: These amino acids are absolutely necessary to produce catalytic activity. Other substitutions are possible, such as deletions of the 10 first amino acids.

FIGURE 3

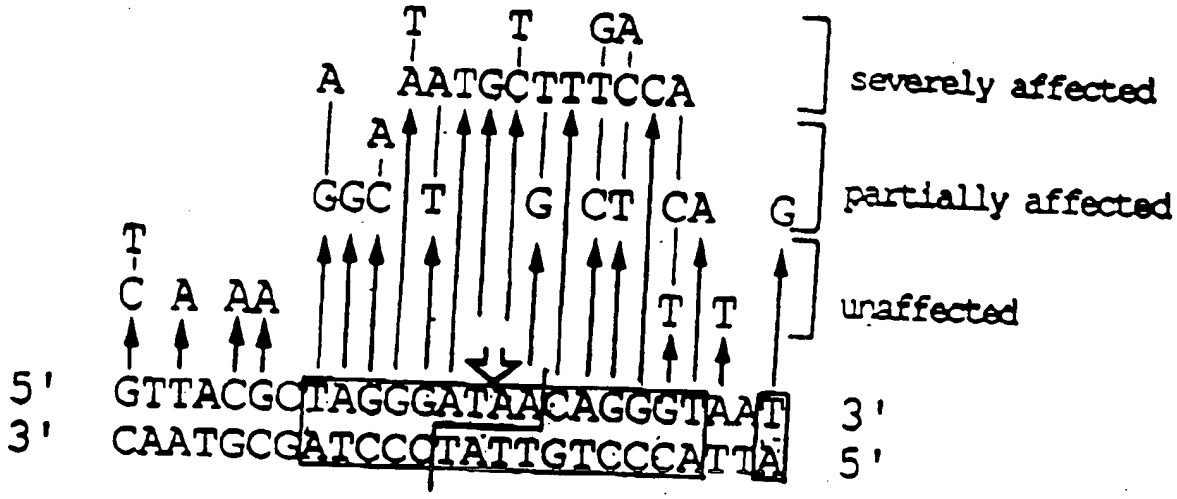
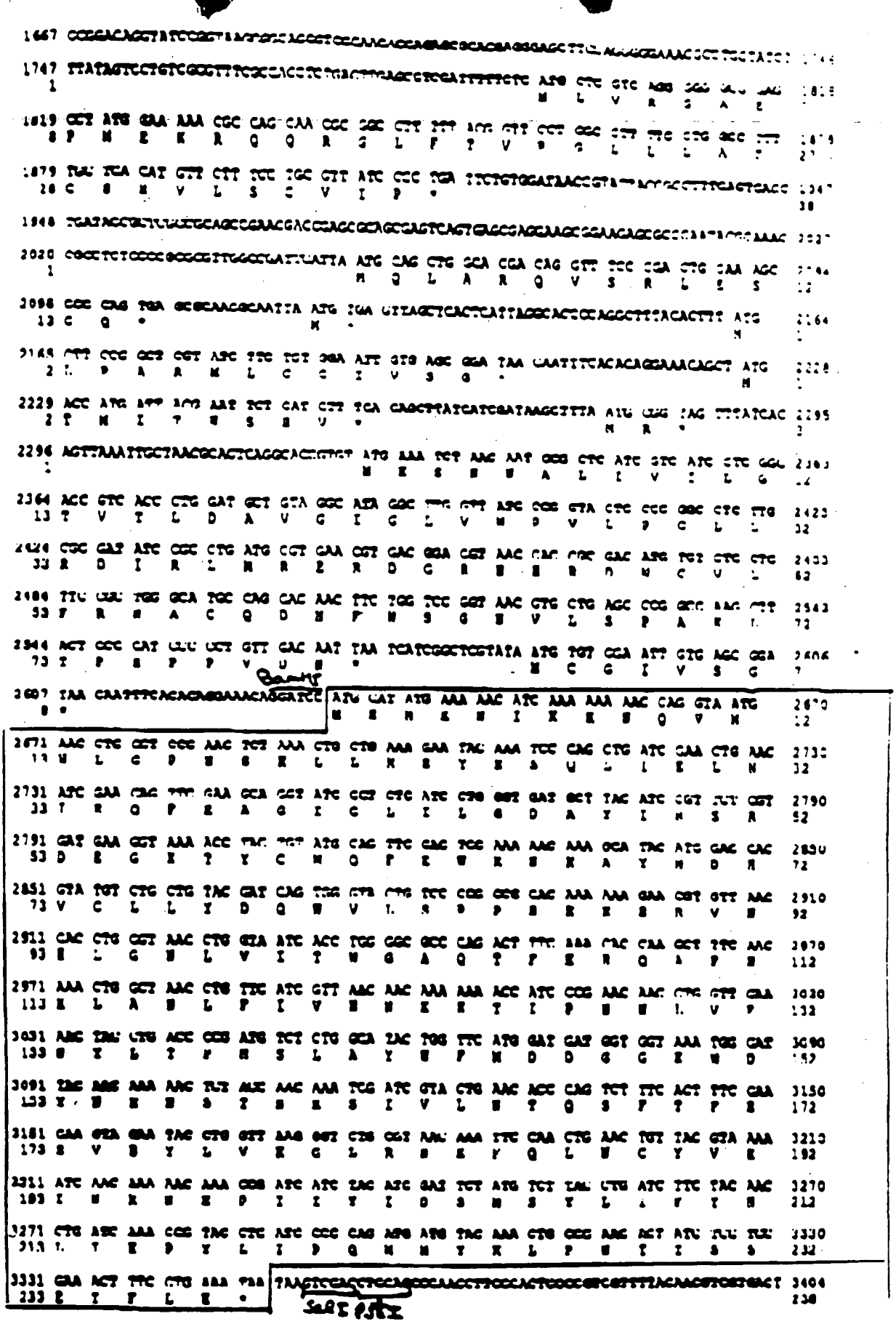


FIGURE 4



I-*Sce*I coding sequence of pSCM525 - Note the two amino acid N-terminal extension as compared to genuine version of the gene.

FIGURE 5

VARIATIONS AROUND THE I-SceI SEQUENCE

	-2	-1	1				5					10					
	M	H	N	K	N	I	K	K	N	Q	V	S	N	L	G	P	N
20												30					
K	L	L	K	E	Y	K	S	Q	L	I	E	L	N	I	E	Q	F
40												50					
G	L	S	L	I	L	G	Q	A	V	I	R	S	R	D	E	G	K
60												70					
C	H	Q	F	E	W	K	N	K	A	Y	M	D	H	V	C	L	_
80												90					
Q	W	V	L	S	P	P	H	K	K	E	R	V	N	M	L	G	N
100												110					
I	T	W	G	A	Q	T	F	K	H	Q	A	F	N	K	L	A	N
120												130					
I	V	N	N	K	K	I	I	P	N	N	L	V	E	N	Y	L	T
140												150					
Q	L	A	Y	W	F	M	Q	Q	Q	Q	K	W	D	Y	N	K	N
160												170					
N	K	S	I	V	L	N	T	Q	S	F	T	F	E	E	V	E	Y
180												190					
K	G	L	R	N	K	F	Q	L	N	C	Y	V	K	I	N	K	N
200												210					
I	I	Y	I	D	G	M	S	Y	L	I	F	Y	N	L	I	K	P
220												230					
I	P	Q	M	M	Y	K	L	P	N	T	I	S	S	E	T	F	L

Positions than can be changed without affecting endonuclease activity (demonstrated)
positions -1 and -2 are not natural. The two amino acids are added due to cloning strategies
positions 1 to 10: can be deleted
position 36: G is tolerated
position 40: M or V are tolerated
position 41: S or N are tolerated
position 43: A is tolerated
position 46: V or N are tolerated
position 91: A is tolerated
positions 138 and 139: L are tolerated
position 222: A and S are tolerated

Changes that affect endonuclease activity (demonstrated)
position 19: L to S
position 38: I to S or N
position 39: G to D or R
position 40: L to Q
position 42: L to R
position 44: D to E, G or H
position 45: A to E or D
position 46: Y to D
position 47: I to R or N
position 80: L to S
position 144: D to E
position 145: D to E
position 146: G to E

FIGURE 6

Group I Intron Encoded Endonucleases and Related Endonucleases

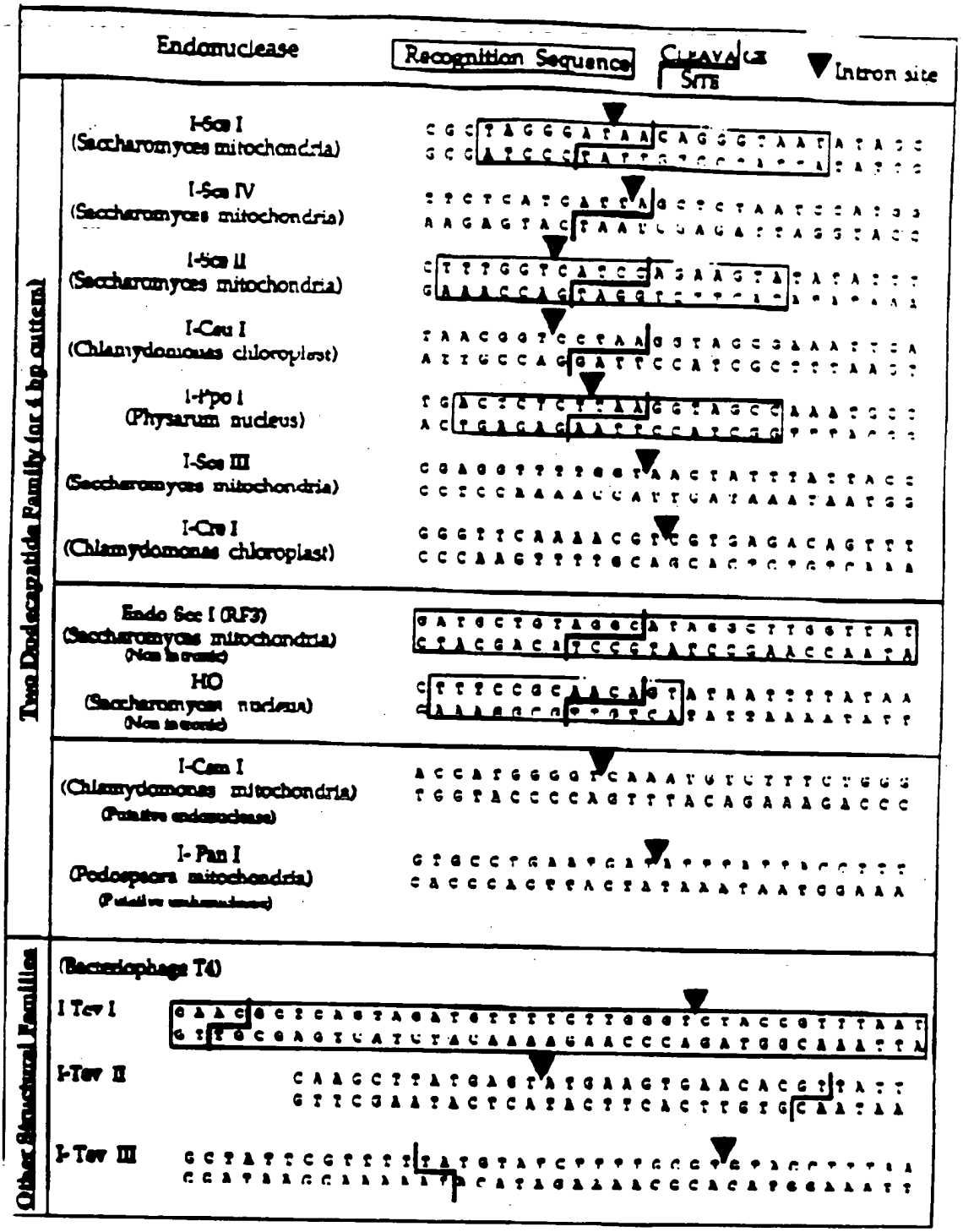


FIGURE 7

EXPRESSION VECTORS

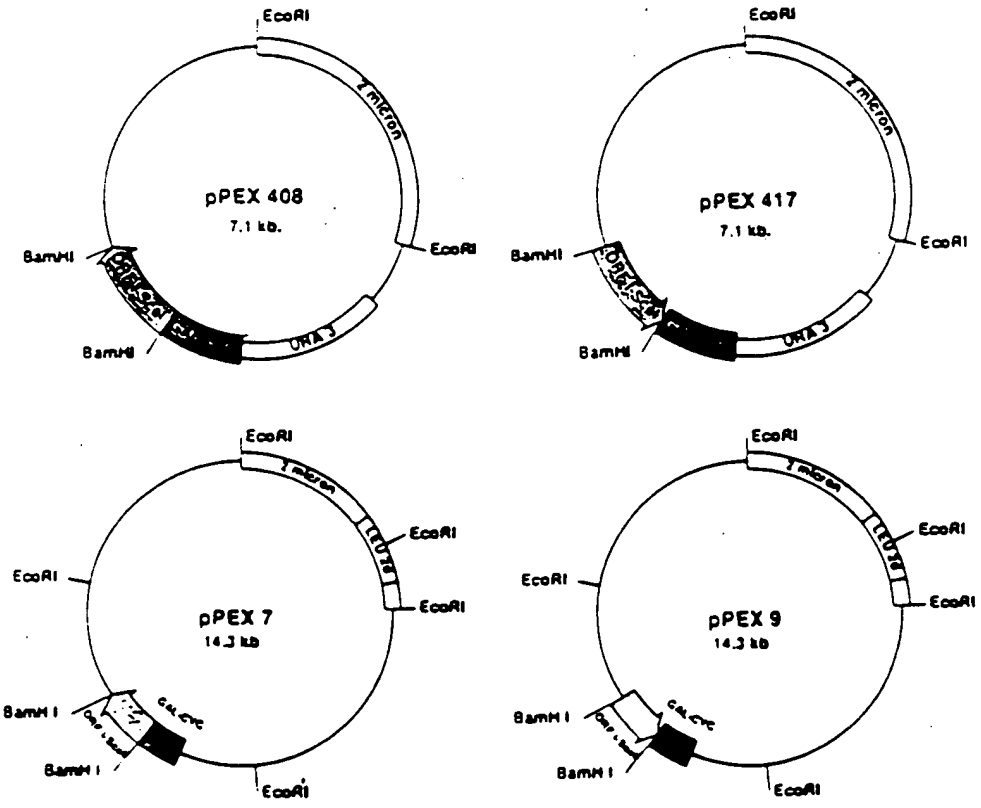
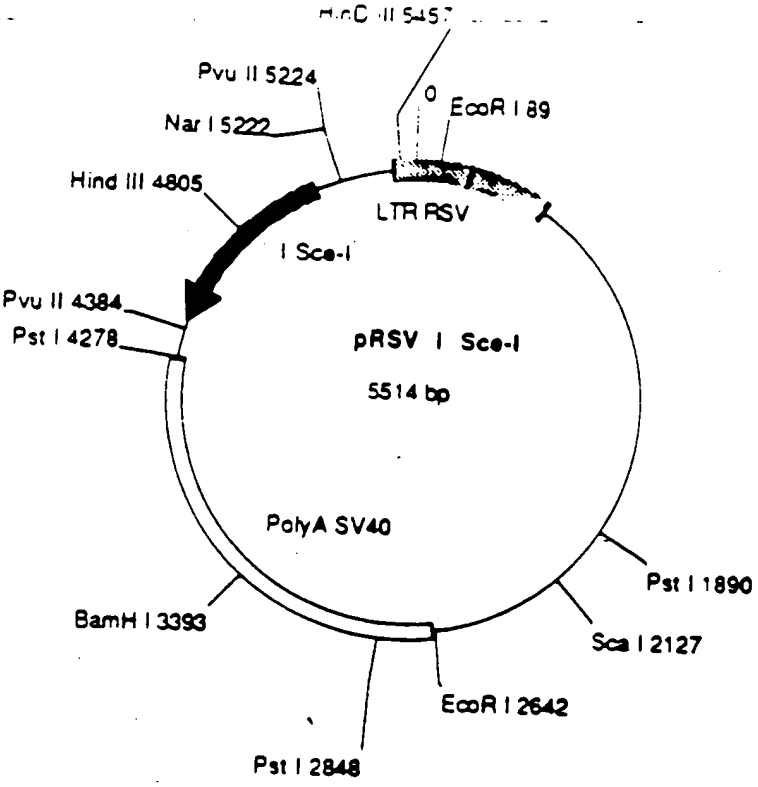


FIGURE 8


Pvu II 5224
Nar I 5222
Hind III 4805
Pvu II 4384
Pst I 4278
BamH I 3393
Pst I 2848
EcoR I 2642
Sca I 2127
Pst I 1890
EcoR I 89
LTR RSV
I Sce-I
pRSV I Sce-I
5514 bp
PolyA SV40

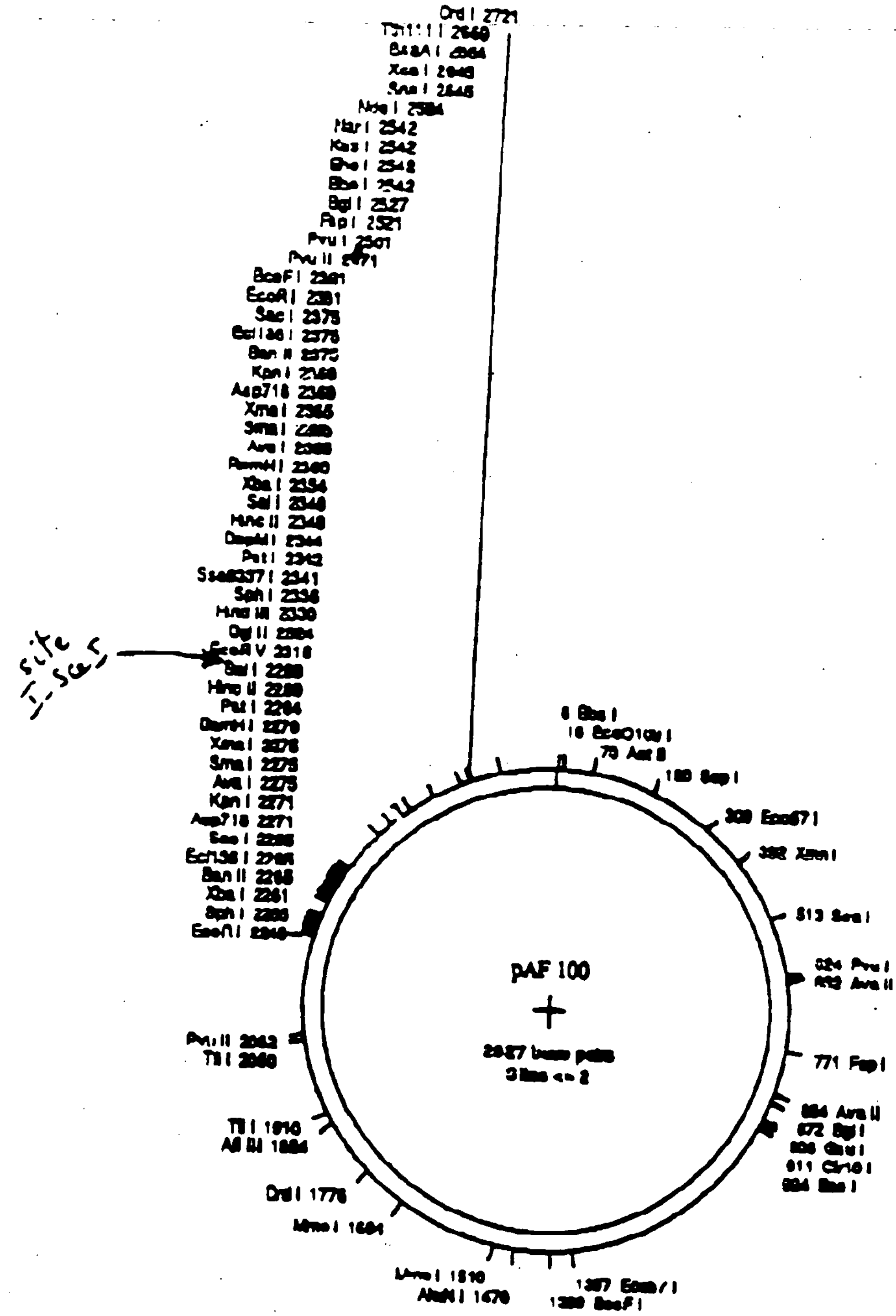

Drd I 2721
Nde I 2584
Kas I 2542
Bgl I 2527
Fsp I 2521
EcoR I 2381
Sac I 2378
Kpn I 2369
Asp718 2369
Xma I 2365
Sma I 2365
Xba I 2354
Sal I 2348
Hinc II 2348
Pst I 2342
Sse8387 I 2341
Sph I 2336
Hind III 2330
EcoR V 2318
site I-SceI
Kpn I 2271
Asp718 2271
Ban II 2265
Xba I 2261
6 Bbs I
79 Aat II
309 Eco57 I
392 Xmn I
771 Fsp I
pAF 100
2927 base pairs
Pvu II 2062
Drd I 1776
AlwN I 1470

FIG. 10A

pAF 100 - RESTRICTION MAP

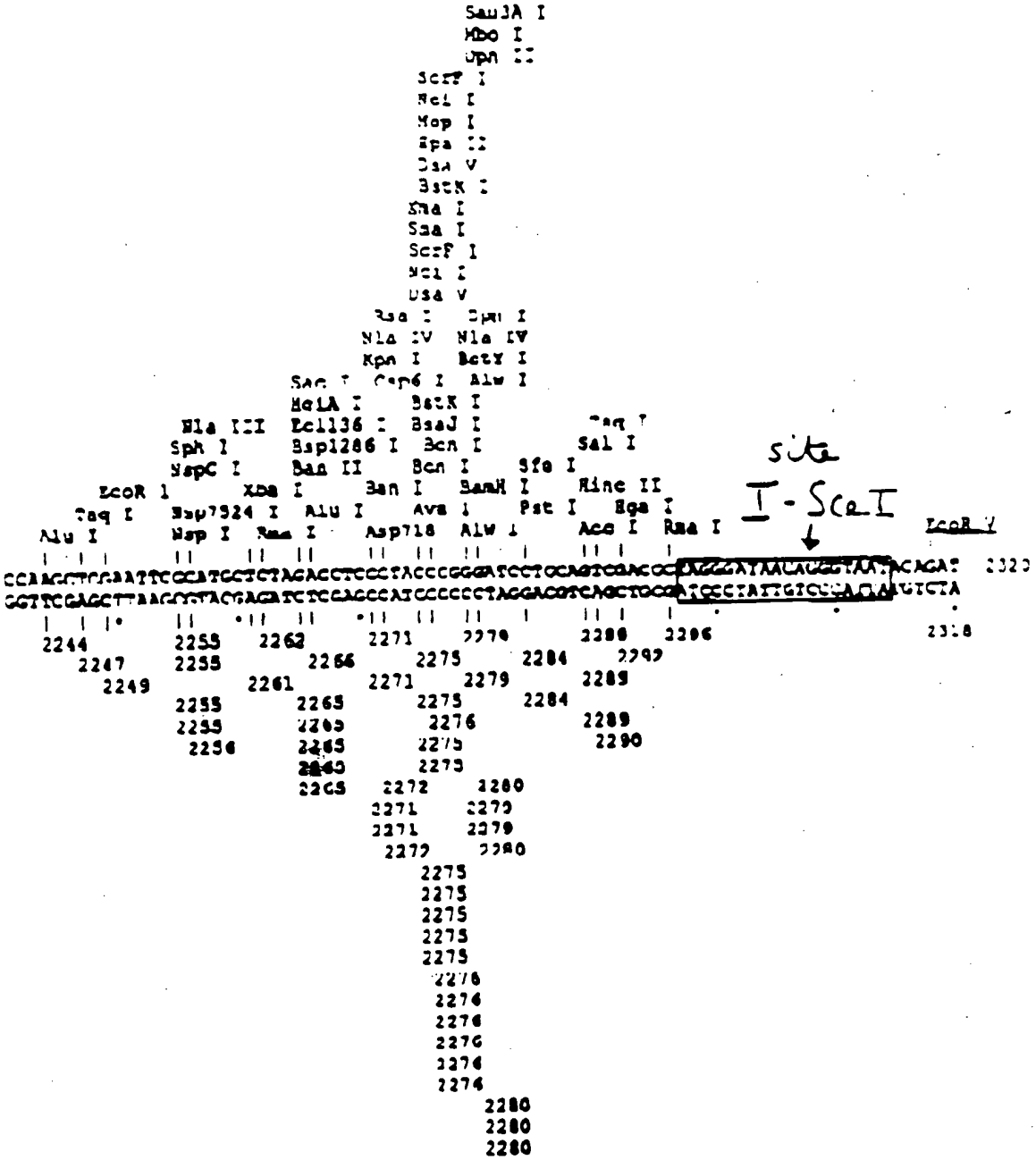


FIGURE 10B

pAF 100 - RESTRICTION MAP

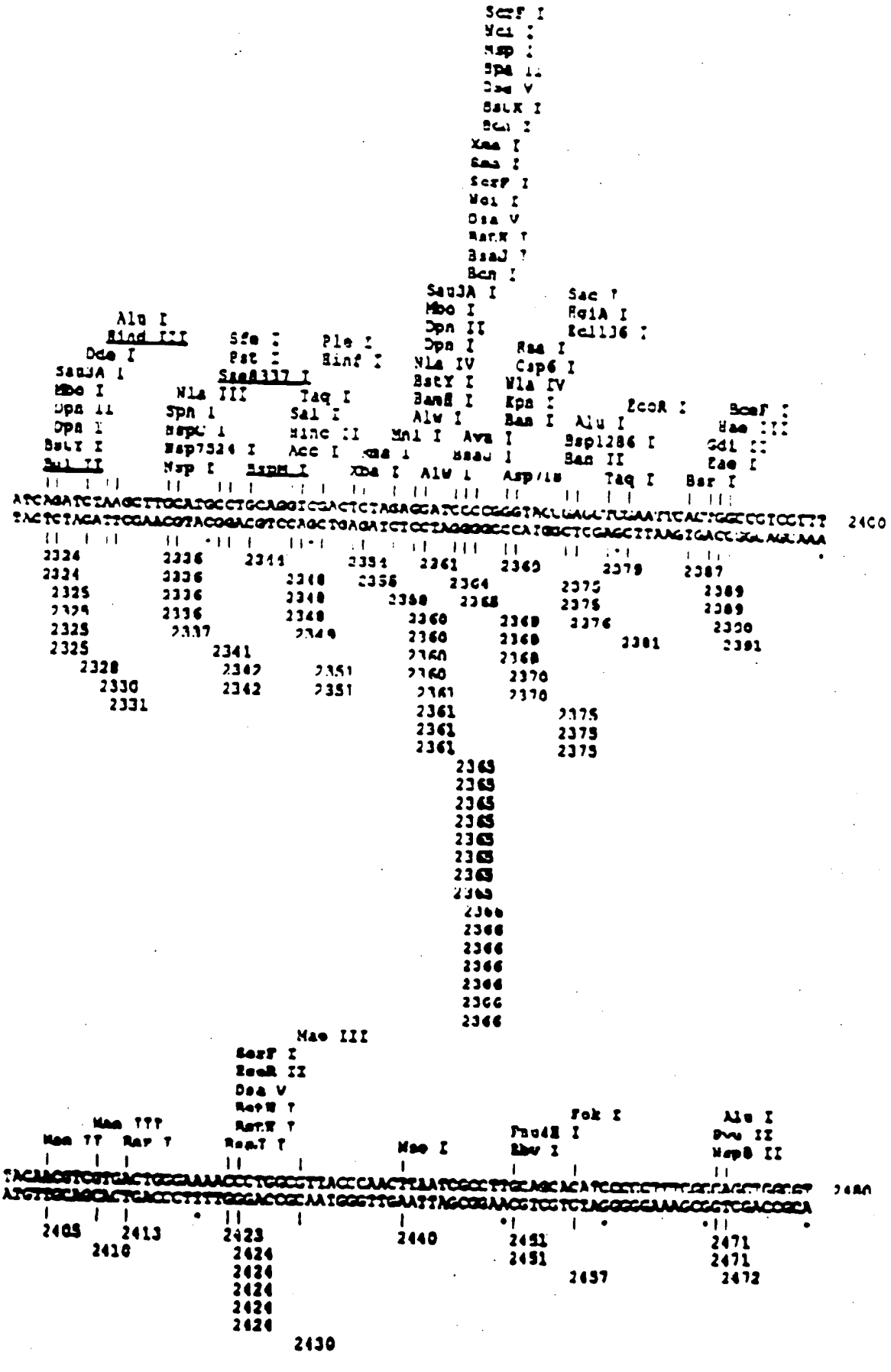
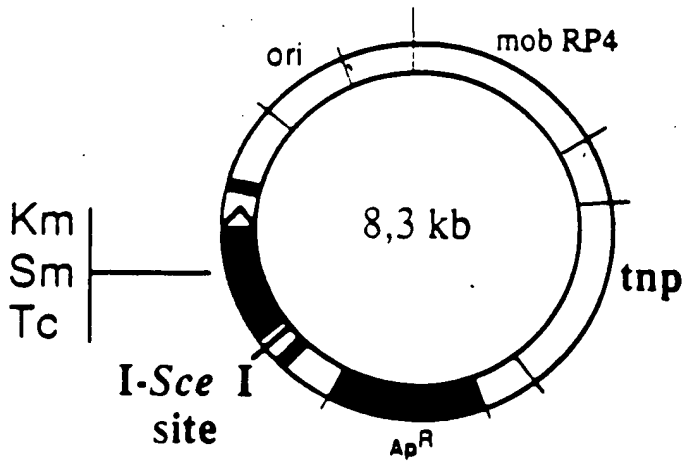


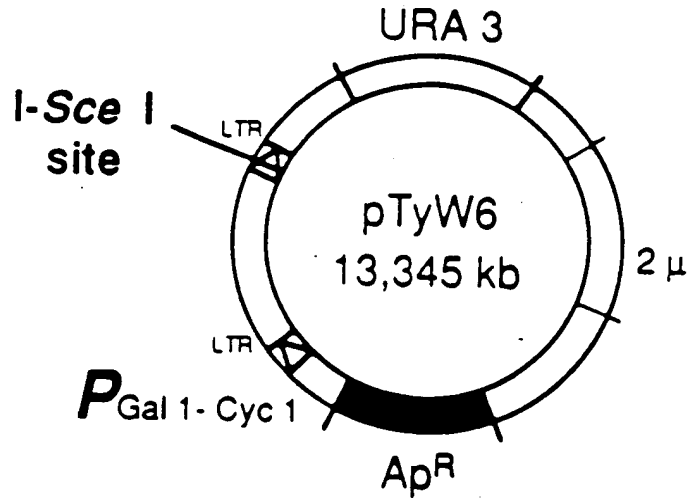
FIGURE 11

Names : pTSm ω
pTKm ω
pTTc ω



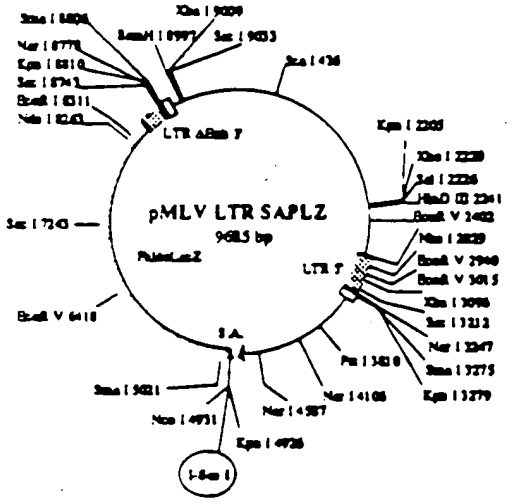
Construction : pGP 704 from De Lorenzo, with transposase gene and insertion of the linker [I-*Sce*I] in NotI unique site

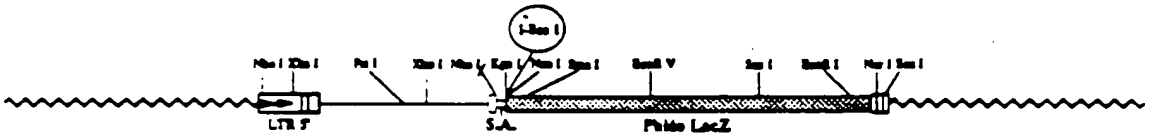
FIGURE 12

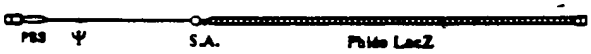


Construction : pD 123 , from J. D. Boeke
with insertion of a linker [I-*Sce*I - NotI] in BamHI

FIGURE 13


pMLV LTR SAPLZ
9685 bp
PhleoLacZ
LTR 5'
S.A.
I-SceI


I-SceI
LTR 5'
S.A.
PhleoLacZ


PBS
Ψ
S.A.
PhleoLacZ

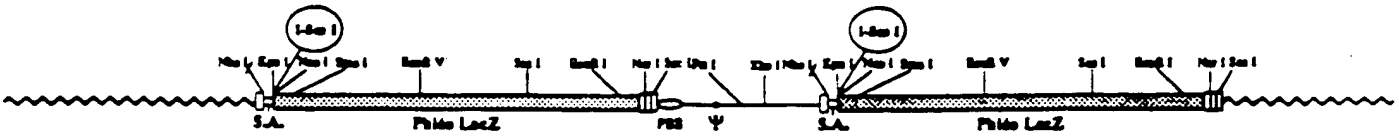

I-SceI
I-SceI
S.A.
PhleoLacZ
PBS
Ψ
S.A.
PhleoLacZ

FIGURE 14

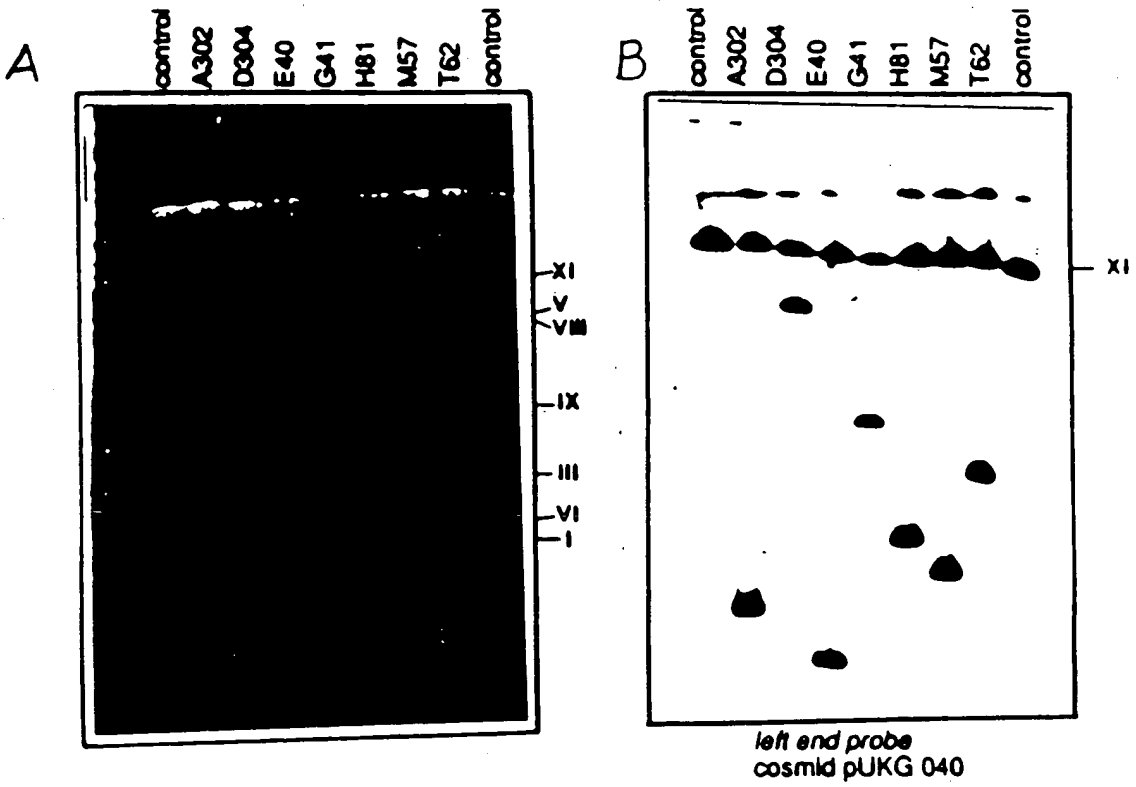

A
control
A302
D304
E40
G41
H81
M57
T62
control
XI
V
VIII
IX
III
VI
I
B
control
A302
D304
E40
G41
H81
M57
T62
control
XI
left end probe
cosmid pUKG 040

FIGURE 15

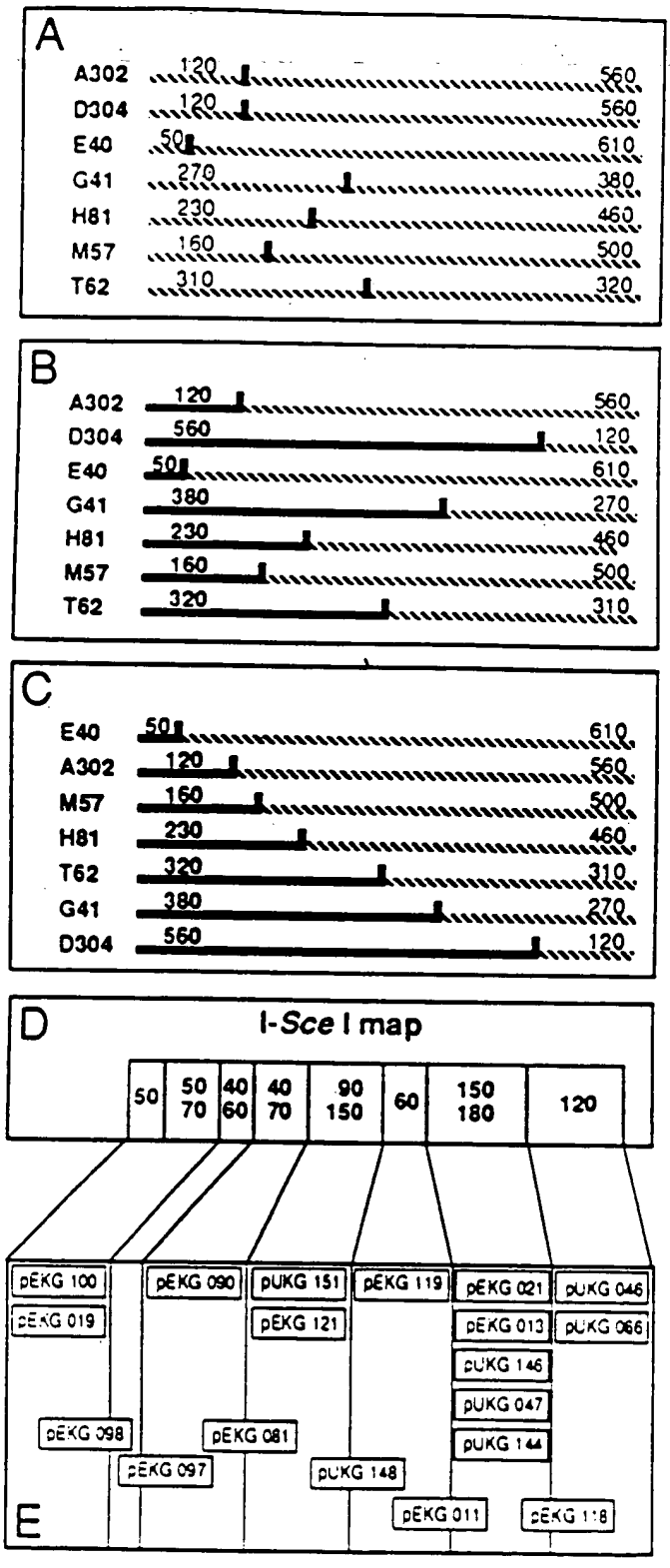

A
A302 120 560
D304 120 560
E40 50 610
G41 270 380
H81 230 460
M57 160 500
T62 310 320
B
A302 120 560
D304 560 120
E40 50 610
G41 380 270
H81 230 460
M57 160 500
T62 320 310
C
E40 50 610
A302 120 560
M57 160 500
H81 230 460
T62 320 310
G41 380 270
D304 560 120
D
I-Sce I map
50
50 70
40 60
40 70
90 150
60
150 180
120
pEKG 100
pEKG 019
pEKG 090
pUKG 151
pEKG 121
pEKG 119
pEKG 021
pUKG 046
pEKG 013
pUKG 066
pUKG 146
pUKG 047
pUKG 144
pEKG 098
pEKG 081
pEKG 097
pUKG 148
pEKG 011
pEKG 118
E

FIGURE 16

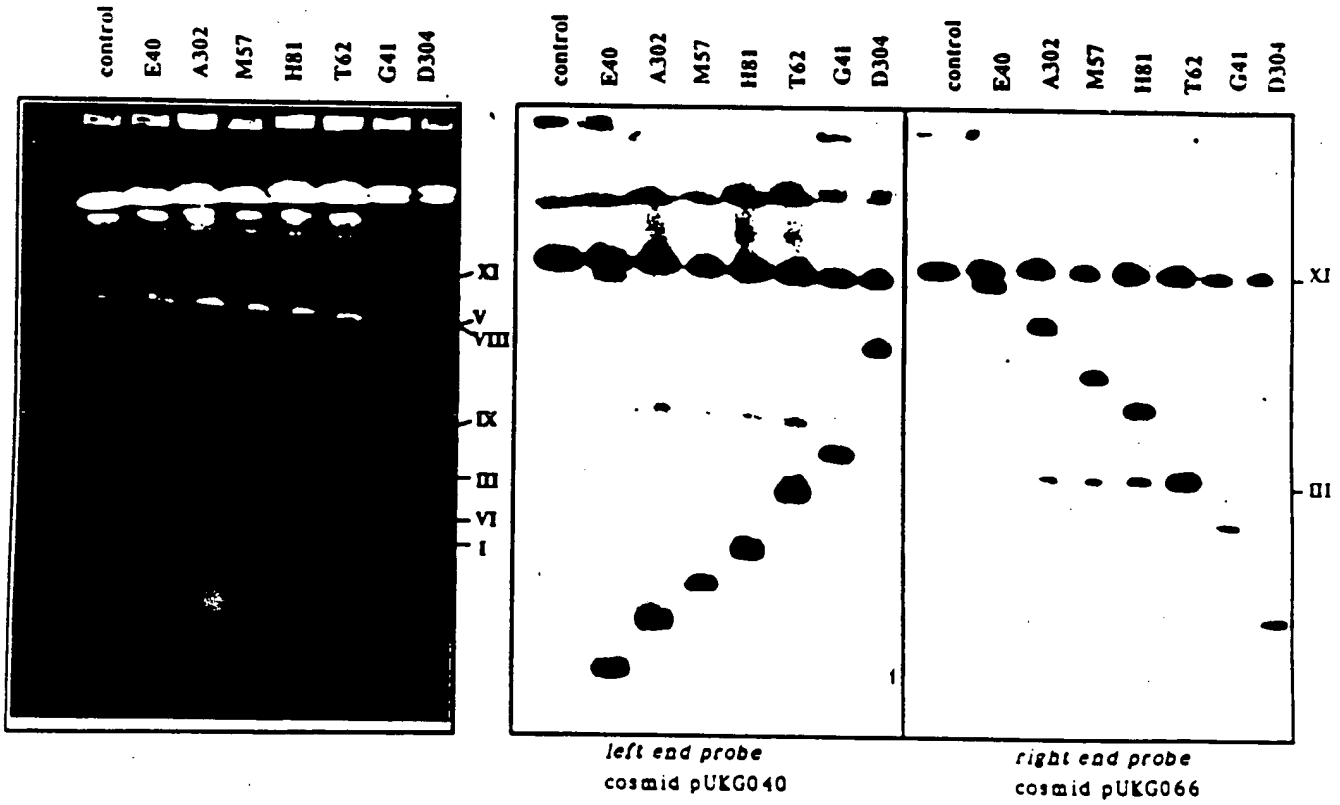

control
E40
A302
M57
H81
T62
G41
D304
XI
V
VIII
IX
III
VI
I
left end probe
cosmid pUKG040
right end probe
cosmid pUKG066

FIGURE 17

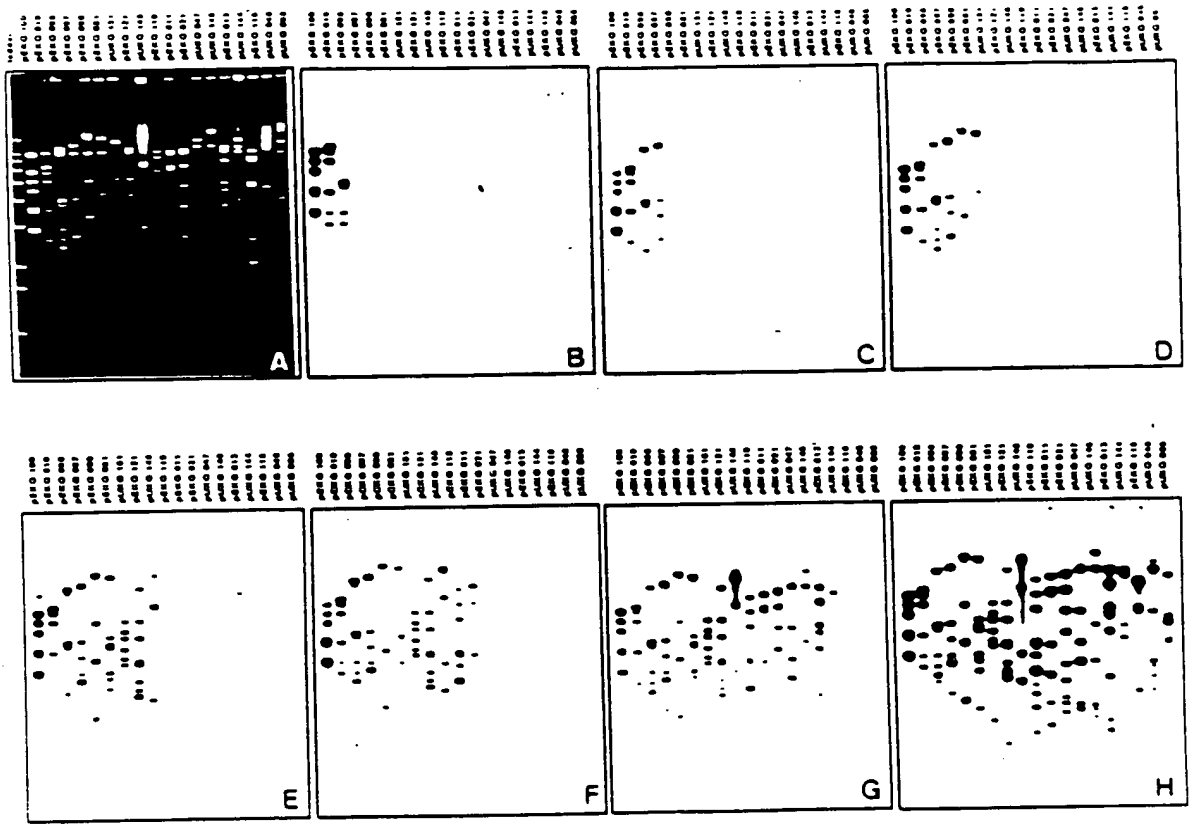


FIGURE 18

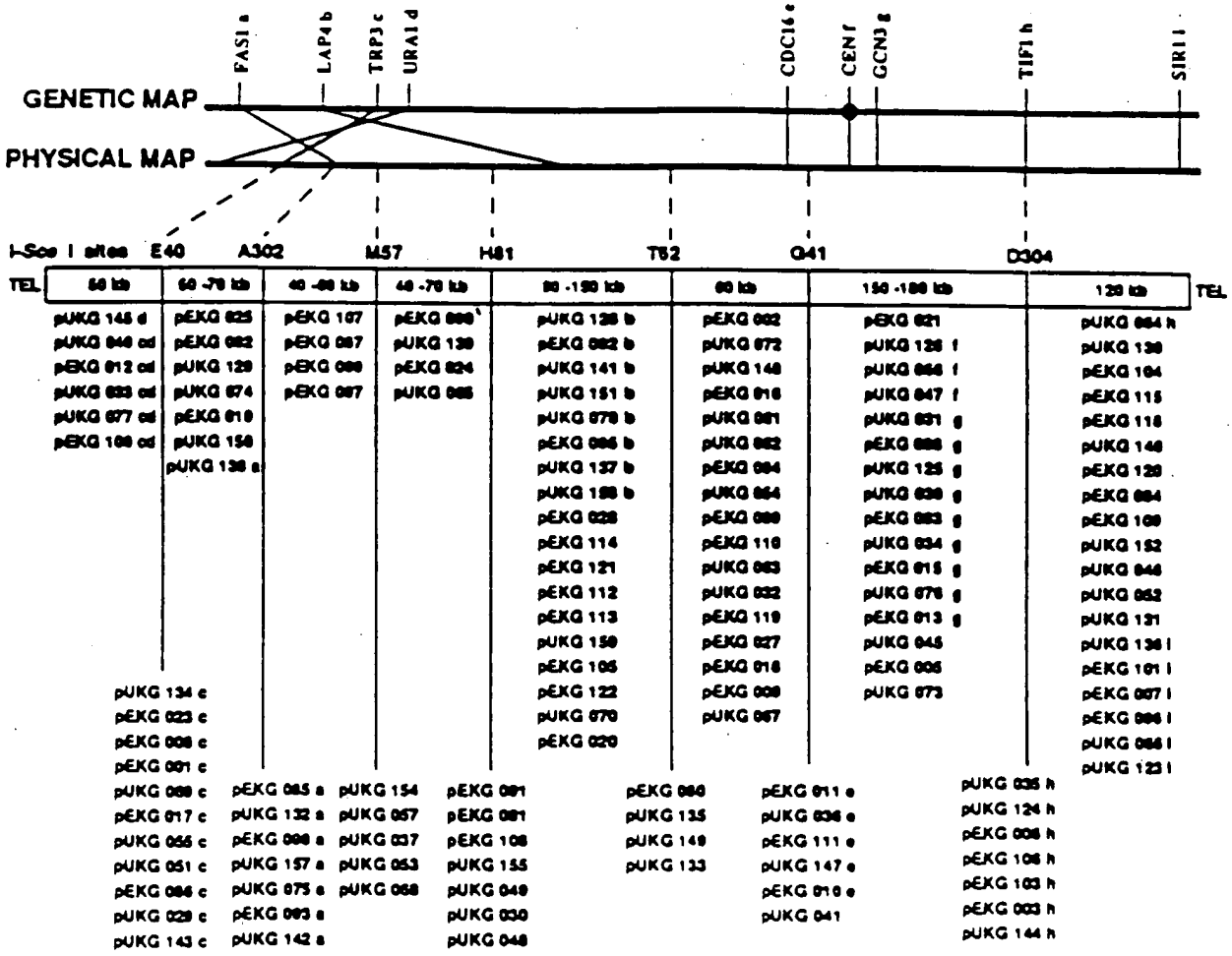
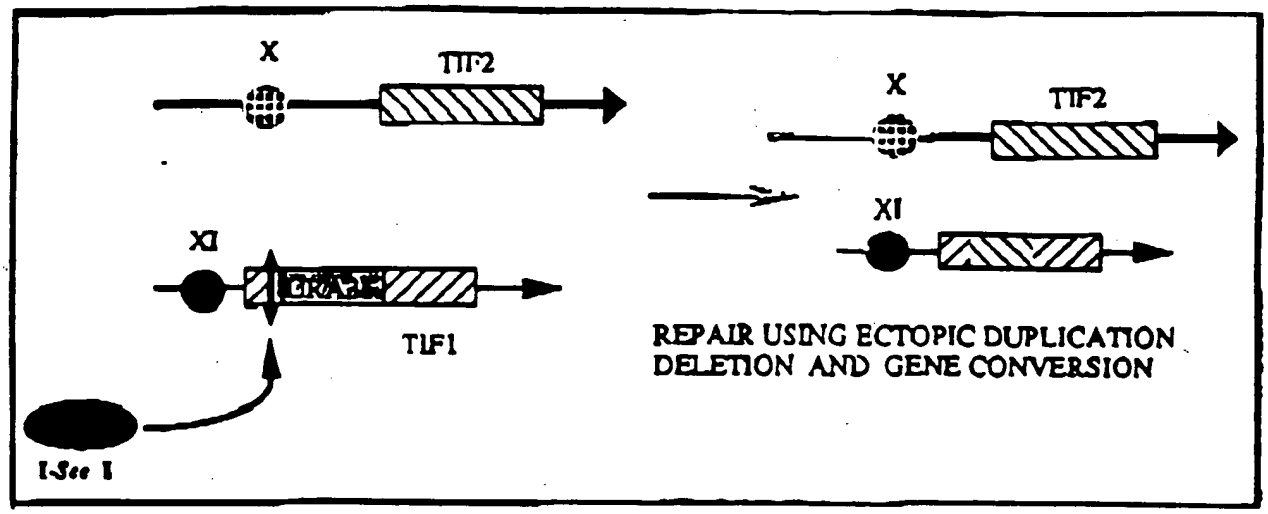

FAS1 a
LAP4 b
TRP1 c
URA1 d
CDC16 e
CEN f
GCN3 g
TIF1 h
SIR1 i
GENETIC MAP
PHYSICAL MAP
I-Sce I sites
E40
A302
M57
H81
T62
G41
D304
TEL
50 kb
50 -70 kb
40 -60 kb
40 -70 kb
90 -150 kb
60 kb
150 -180 kb
120 kb
TEL
pUKG 145 d
pUKG 040 cd
pEKG 012 cd
pUKG 033 cd
pUKG 077 cd
pEKG 100 cd
pEKG 025
pEKG 082
pUKG 129
pUKG 074
pEKG 019
pUKG 150
pUKG 136 a
pEKG 107
pEKG 087
pEKG 099
pEKG 097
pUKG 099
pUKG 139
pEKG 024
pUKG 088
pUKG 126 b
pEKG 082 b
pUKG 141 b
pUKG 151 b
pUKG 078 b
pEKG 096 b
pUKG 137 b
pUKG 155 b
pEKG 028
pEKG 114
pEKG 121
pEKG 112
pEKG 113
pUKG 150
pEKG 105
pEKG 122
pUKG 070
pEKG 020
pEKG 002
pUKG 072
pUKG 148
pEKG 016
pUKG 081
pUKG 082
pEKG 004
pUKG 064
pEKG 099
pEKG 110
pUKG 083
pUKG 032
pEKG 119
pEKG 027
pEKG 018
pEKG 009
pUKG 067
pEKG 021
pUKG 126 f
pUKG 066 f
pUKG 047 f
pUKG 031 g
pEKG 098 g
pUKG 125 g
pUKG 038 g
pEKG 083 g
pUKG 034 g
pEKG 015 g
pUKG 076 g
pEKG 013 g
pUKG 045
pEKG 005
pUKG 073
pUKG 064 h
pUKG 139
pEKG 104
pEKG 115
pEKG 116
pUKG 146
pEKG 129
pEKG 004
pEKG 109
pUKG 152
pUKG 046
pUKG 052
pUKG 131
pUKG 136 i
pEKG 101 i
pEKG 007 i
pEKG 006 i
pUKG 064 i
pUKG 123 i
pUKG 134 c
pEKG 023 c
pEKG 008 c
pEKG 001 c
pUKG 060 c
pEKG 017 c
pUKG 056 c
pUKG 051 c
pEKG 004 c
pUKG 029 c
pUKG 143 c
pEKG 085 a
pUKG 132 a
pEKG 098 a
pUKG 157 a
pUKG 075 a
pEKG 093 a
pUKG 142 a
pUKG 154
pUKG 057
pUKG 037
pUKG 053
pUKG 058
pEKG 091
pEKG 091
pEKG 108
pUKG 155
pUKG 049
pUKG 030
pUKG 048
pEKG 080
pUKG 135
pUKG 149
pUKG 133
pEKG 011 e
pUKG 036 e
pEKG 111 e
pUKG 147 e
pEKG 010 e
pUKG 041
pUKG 035 h
pUKG 124 h
pEKG 008 h
pEKG 106 h
pEKG 103 h
pEKG 003 h
pUKG 144 h

FIGURE 19


X
TIF2
XI
TIF1
I-Sce I
X
TIF2
XI
REPAIR USING ECTOPIC DUPLICATION
DELETION AND GENE CONVERSION

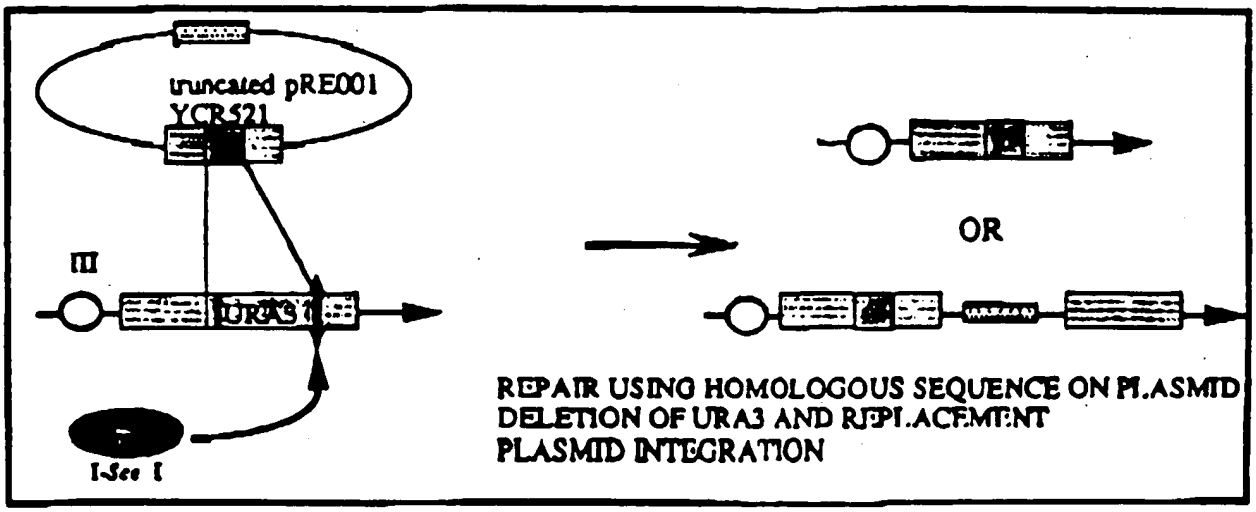

truncated pRE001
YCR521
III
URA3
I-Sce I
OR
REPAIR USING HOMOLOGOUS SEQUENCE ON PLASMID
DELETION OF URA3 AND REPLACEMENT
PLASMID INTEGRATION

Figure 20

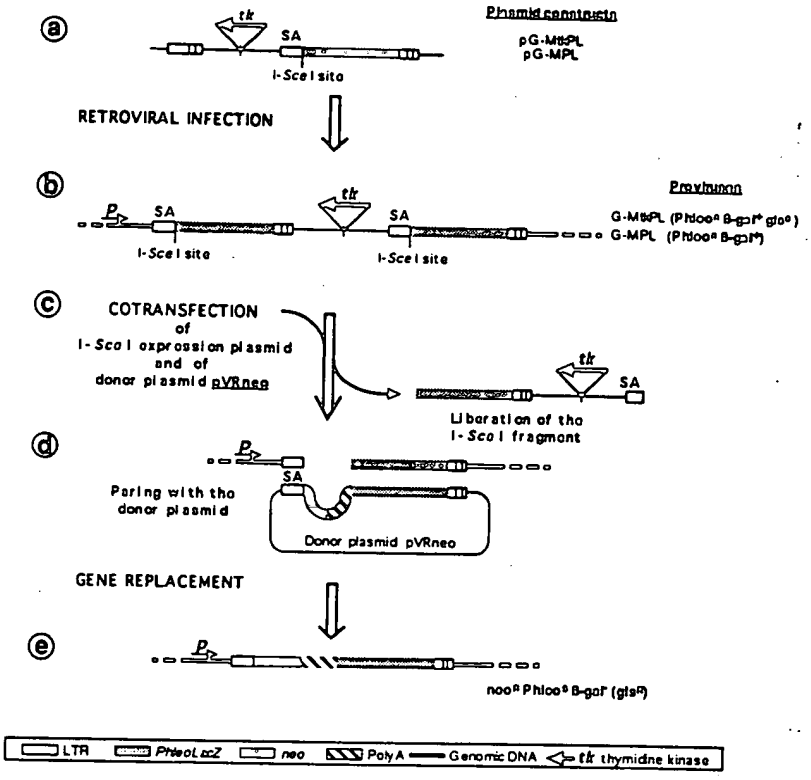

a
tk
SA
I-SceI site
Plasmid constructs
pG-MtiSPL
pG-MPL
RETROVIRAL INFECTION
b
Provirus
P
SA
tk
SA
I-SceI site
I-SceI site
G-MtiSPL (Phleo^R β-gal^+ gls^S)
G-MPL (Phleo^R β-gal^+)
c
COTRANSFECTION
of
I-SceI expression plasmid
and of
donor plasmid pVRneo
tk
SA
Liberation of the
I-SceI fragment
d
P
SA
Pairing with the
donor plasmid
Donor plasmid pVRneo
GENE REPLACEMENT
e
P
neo^R Phleo^S β-gal^+ (gls^R)
LTR
Phleo LacZ
neo
Poly A
Genomic DNA
tk thymidine kinase

Figure 21

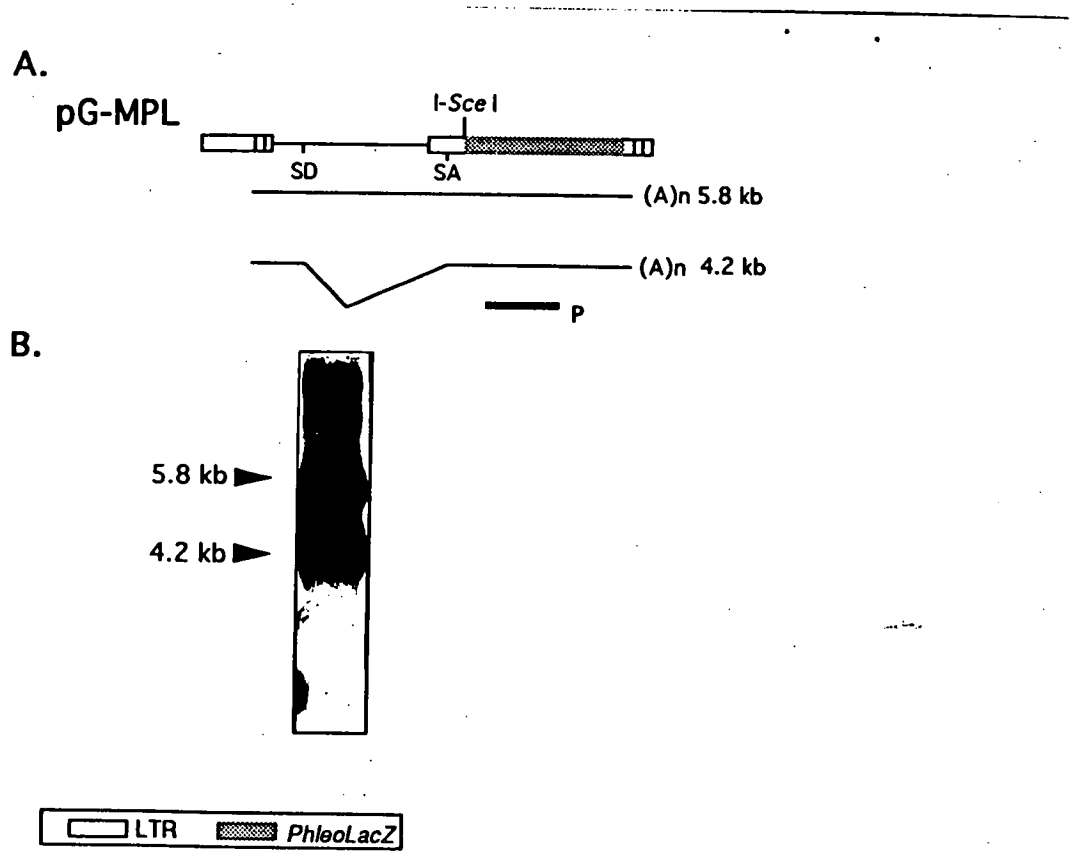


Figure 22

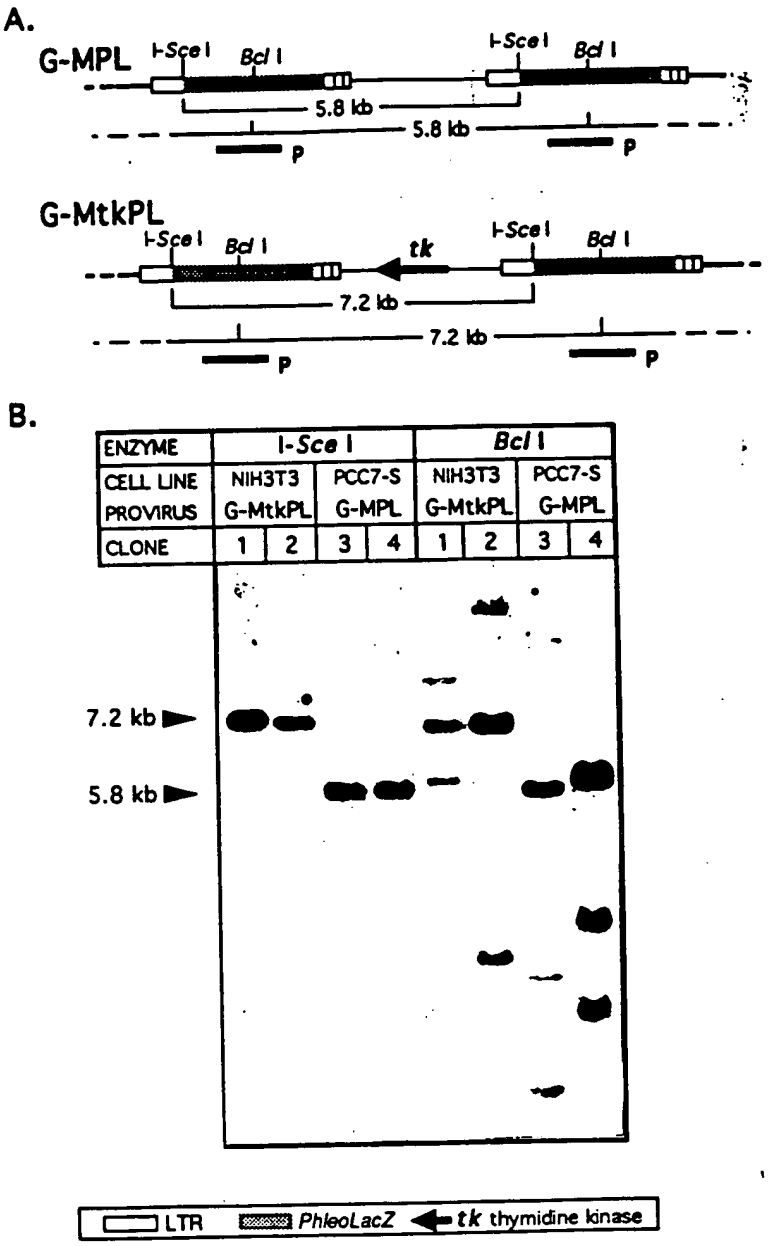

A.
G-MPL
I-Sce I
Bcl I
I-Sce I
Bcl I
5.8 kb
5.8 kb
P
P
G-MtkPL
I-Sce I
Bcl I
tk
I-Sce I
Bcl I
7.2 kb
7.2 kb
P
P
B.
ENZYME
I-Sce I
Bcl I
CELL LINE
NIH3T3
PCC7-S
NIH3T3
PCC7-S
PROVIRUS
G-MtkPL
G-MPL
G-MtkPL
G-MPL
CLONE
1
2
3
4
1
2
3
4
7.2 kb
5.8 kb
LTR
PhleoLacZ
tk thymidine kinase

Figure 23

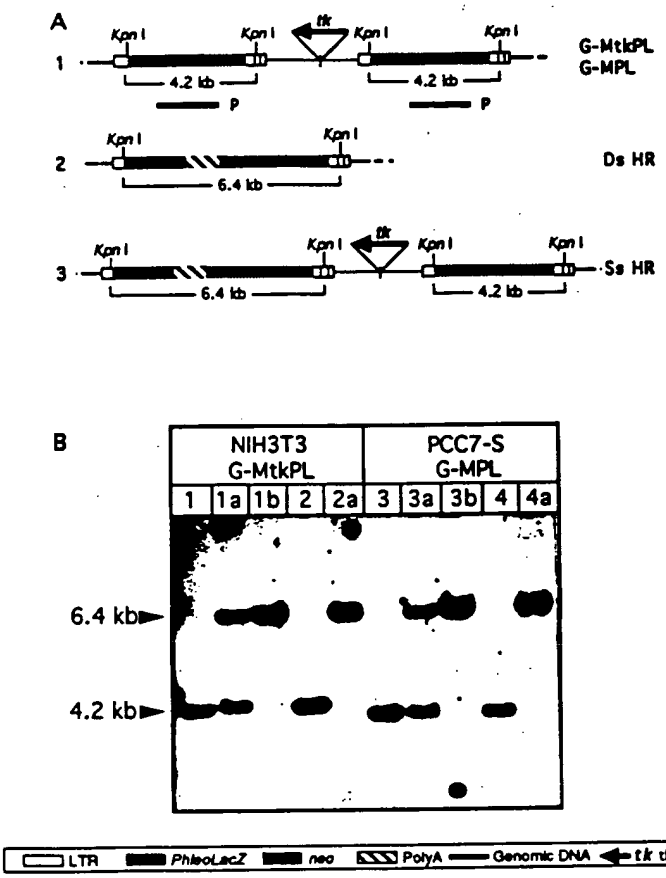
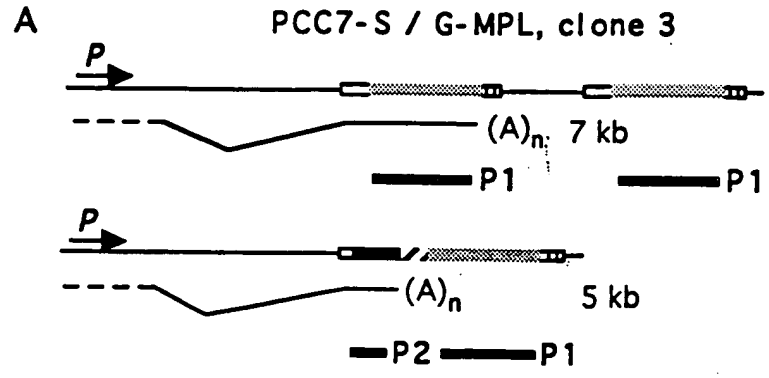
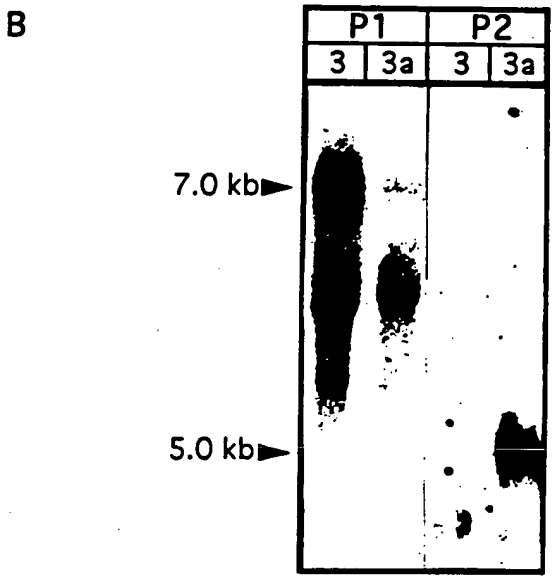

A
Kpn I
tk
1
4.2 kb
P
G-MtkPL
G-MPL
2
6.4 kb
Ds HR
3
Ss HR
B
NIH3T3
G-MtkPL
PCC7-S
G-MPL
1
1a
1b
2
2a
3
3a
3b
4
4a
6.4 kb
4.2 kb
LTR
PhleoLacZ
neo
PolyA
Genomic DNA
tk thymidine kinase

Figure 24


A
PCC7-S / G-MPL, clone 3
P
(A)n 7 kb
P1
P1
P
(A)n
5 kb
P2
P1


B
P1
P2
3
3a
3
3a
7.0 kb
5.0 kb


LTR
PhleoLacZ
neo
PolyA
tk thymidine kinase

Figure 25

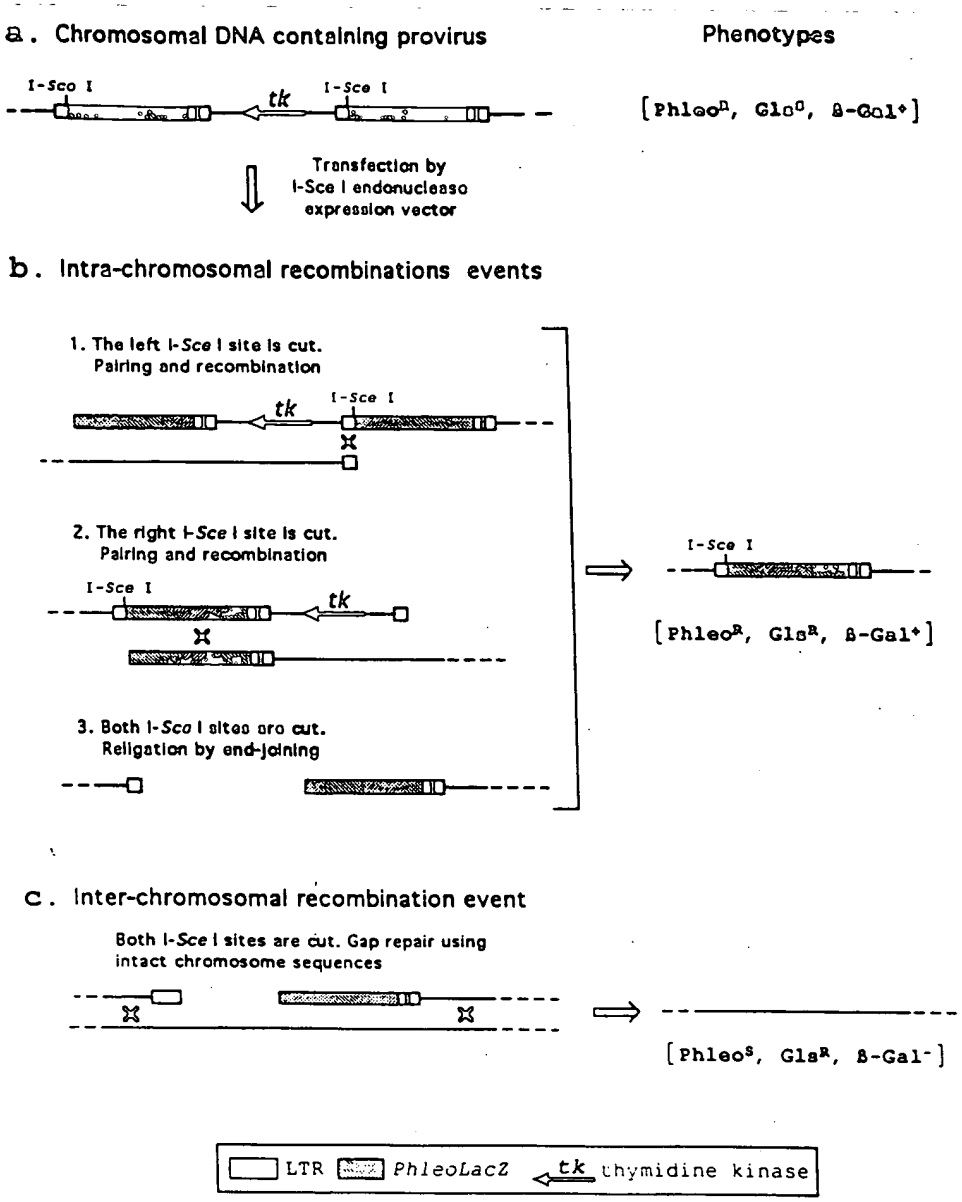
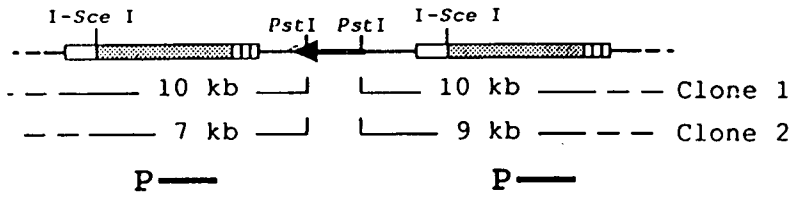
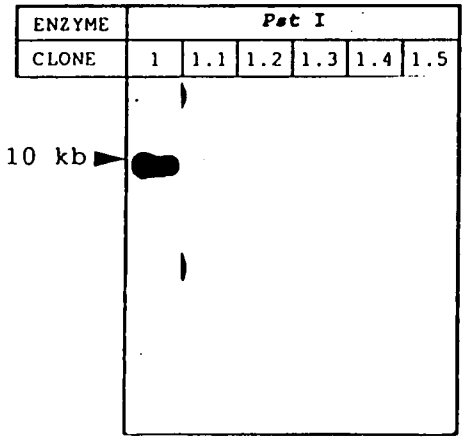

a. Chromosomal DNA containing provirus
Phenotypes
I-Sce I
tk
I-Sce I
[Pheo^R, Gls^S, ß-Gal^+]
Transfection by
I-Sce I endonuclease
expression vector
b. Intra-chromosomal recombinations events
1. The left I-Sce I site is cut.
Pairing and recombination
tk
I-Sce I
2. The right I-Sce I site is cut.
Pairing and recombination
I-Sce I
tk
3. Both I-Sce I sites are cut.
Religation by end-joining
I-Sce I
[Pheo^R, Gls^R, ß-Gal^+]
c. Inter-chromosomal recombination event
Both I-Sce I sites are cut. Gap repair using
intact chromosome sequences
[Pheo^S, Gls^R, ß-Gal^-]
LTR
PheoLacZ
tk thymidine kinase

Figure 26

a. Parental DNA, G-MtkPL



b.



c.

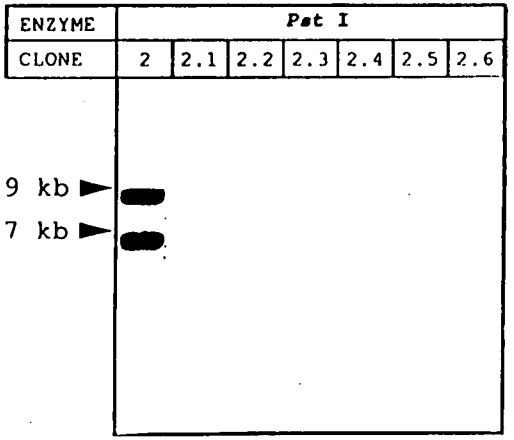


Figure 27

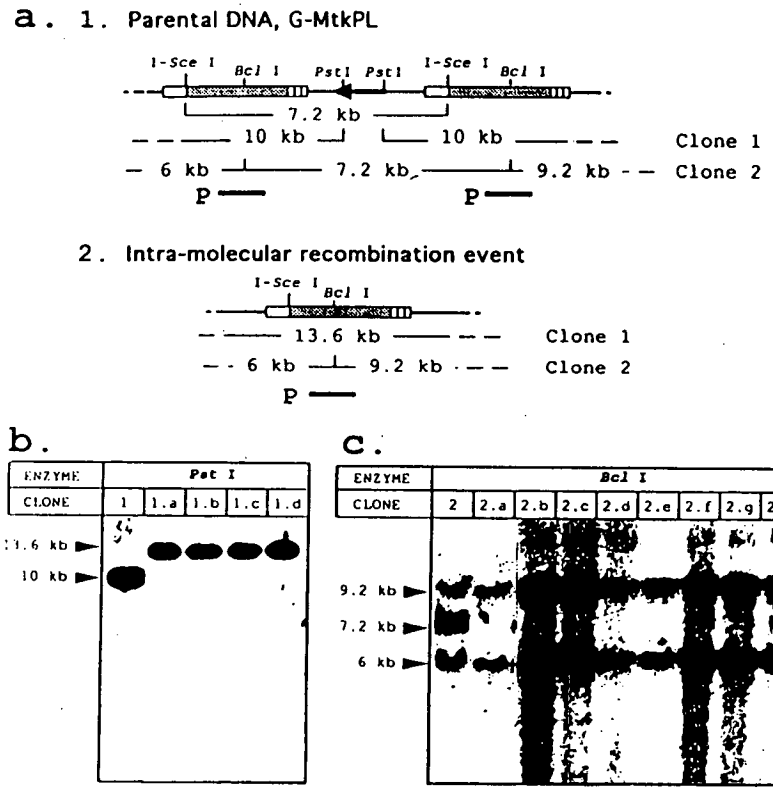
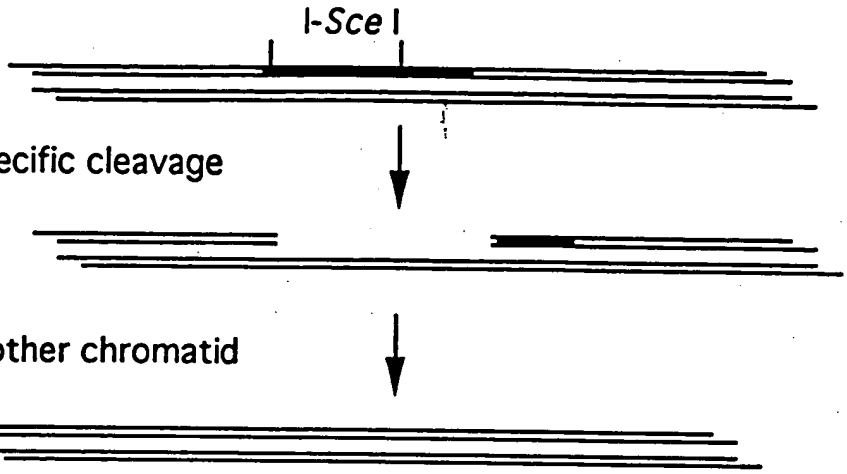

a. 1. Parental DNA, G-MtkPL
I-Sce I
Bcl I
PstI
PstI
I-Sce I
Bcl I
7.2 kb
10 kb
10 kb
Clone 1
6 kb
7.2 kb
9.2 kb
Clone 2
P
P
2. Intra-molecular recombination event
I-Sce I
Bcl I
13.6 kb
Clone 1
6 kb
9.2 kb
Clone 2
P
b.
ENZYME
Pst I
CLONE
1
1.a
1.b
1.c
1.d
13.6 kb
10 kb
c.
ENZYME
Bcl I
CLONE
2
2.a
2.b
2.c
2.d
2.e
2.f
2.g
2.h
9.2 kb
7.2 kb
6 kb

Figure 28

LOSS OF HETEROZYGOSITY

Integration of artificial site or presence of specific site

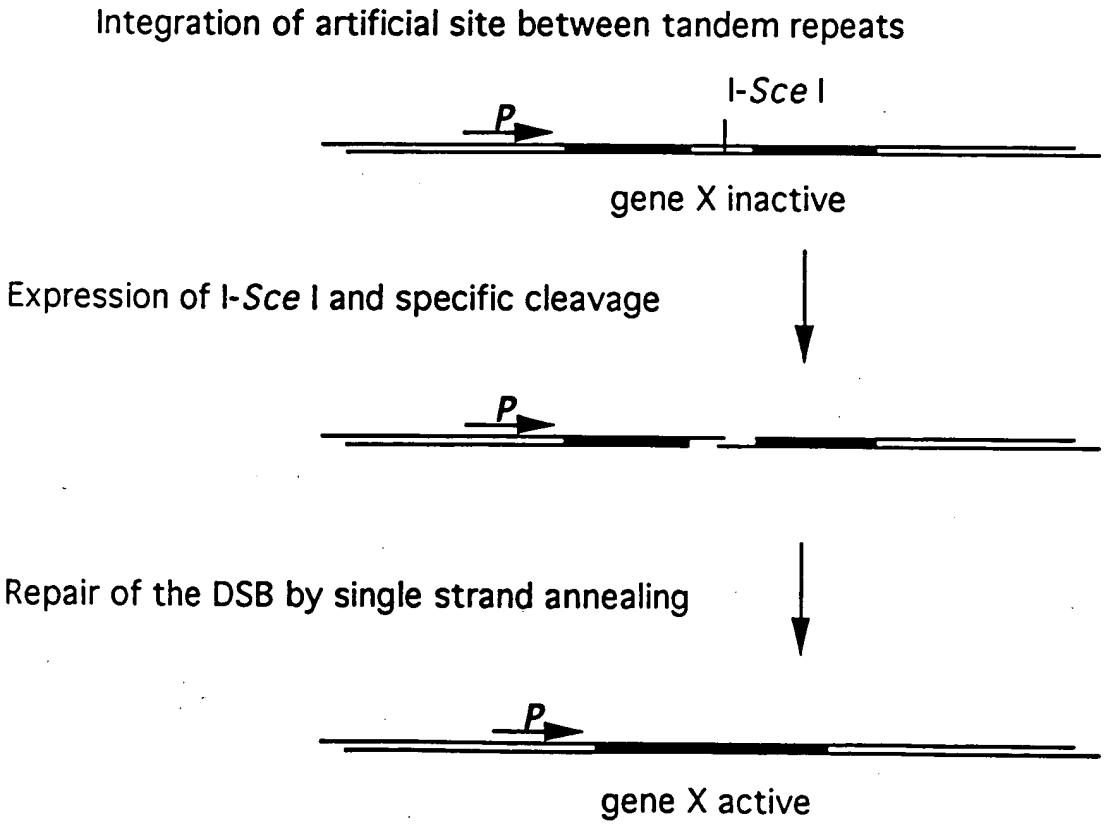


Expression of I-*Sce* I and specific cleavage

Repair of the DSB with the other chromatid

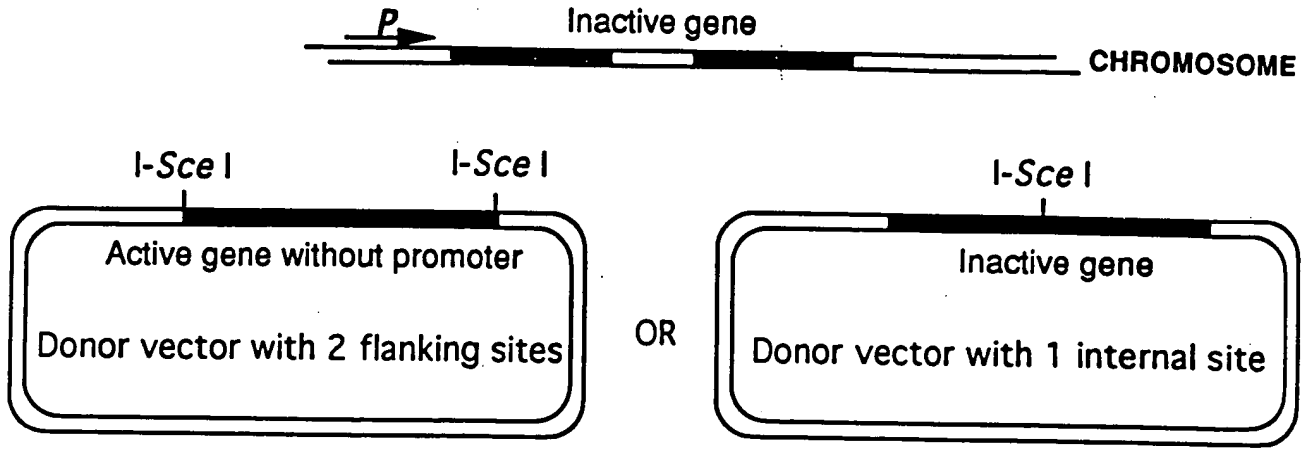
Figure 29

CONDITIONNAL ACTIVATION (Tandem repeat)

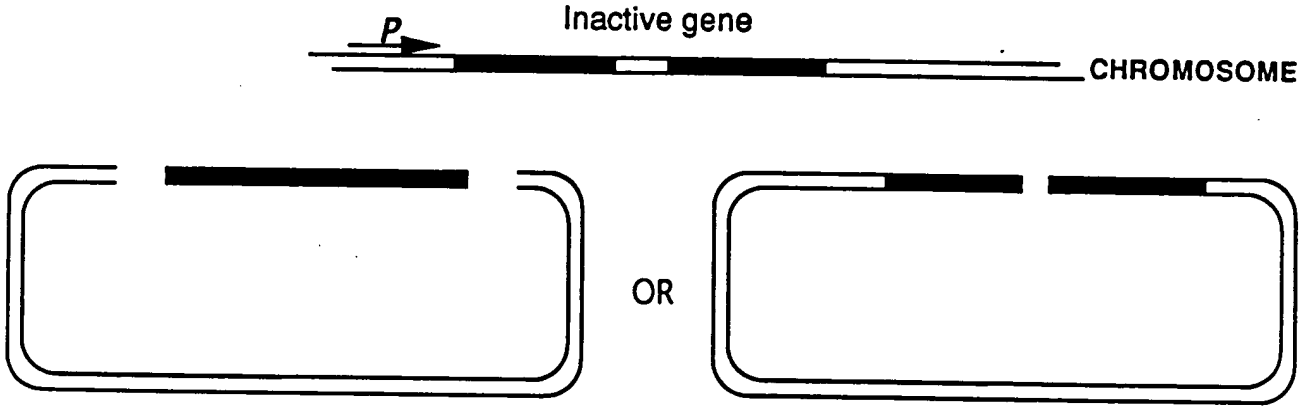


ONE STEP REARRANGEMENT

Integration of artificial site or presence of specific site



Expression of I-*Sce* I enzyme
and
specific cleavage of the donor plasmid



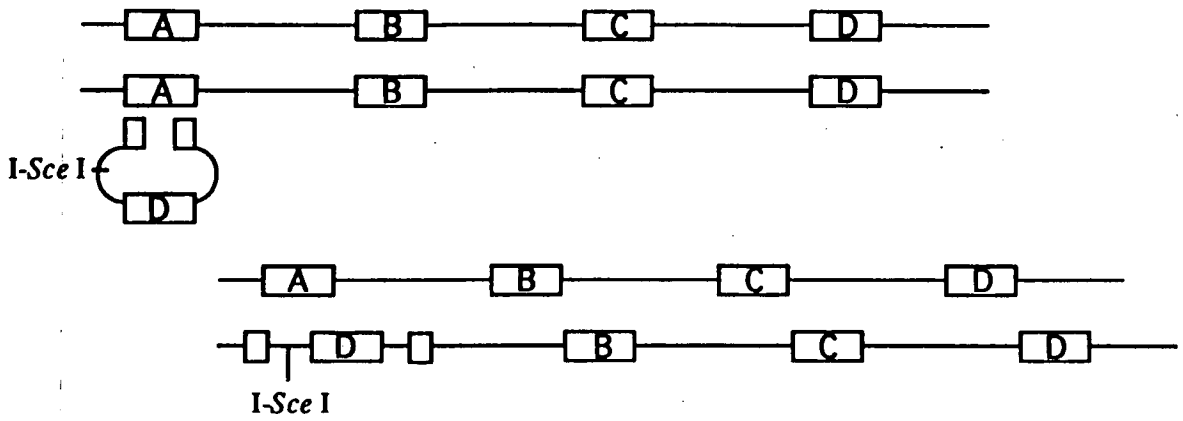
Recombination between the chromosome and plasmid

Active gene

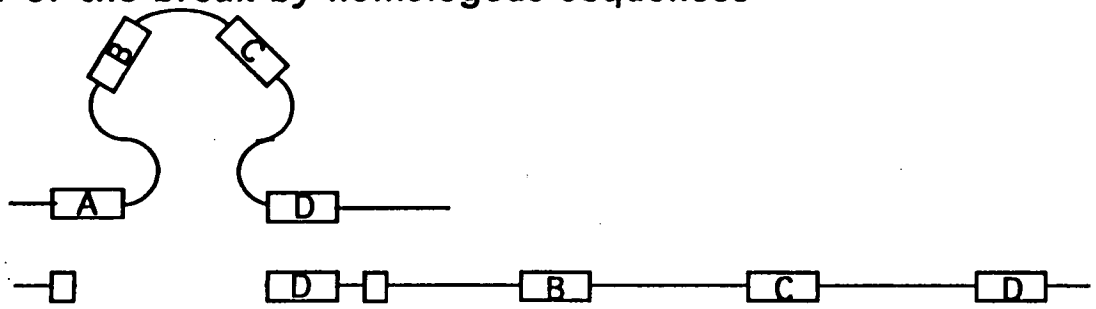
Figure 31

DUPLICATION OF A LOCUS

1 insertion of I-*Sce* I site by classical gene replacement



2 Specific cleavage by I-*Sce* I enzyme
and
repair of the break by homologous sequences



3 Duplication of the totality of the locus

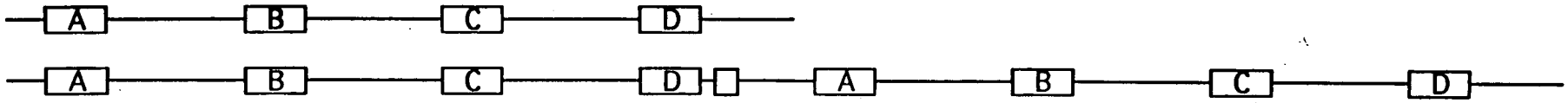
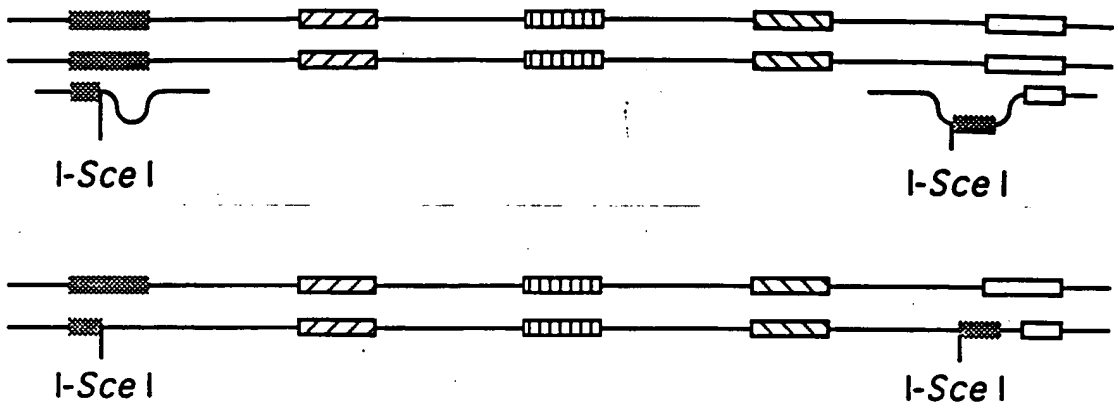


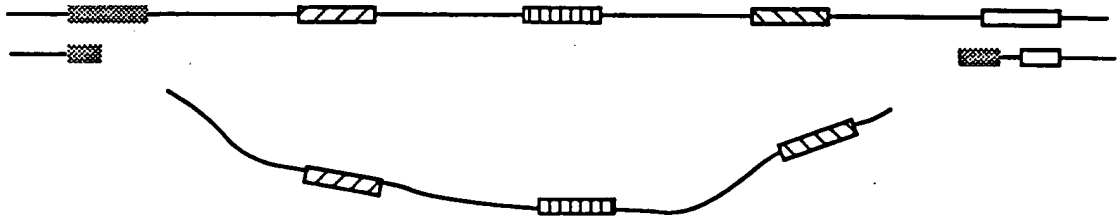
Figure 32

DELETION OF A LOCUS

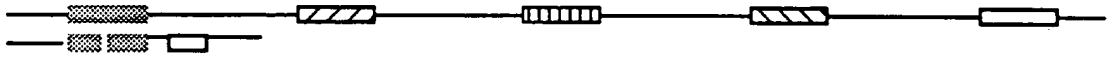
1 Insertion of two I-*Sce* I sites flanking the locus



2 Expression of the enzyme and cleavage



3 Recombination between the two ends



4 deletion of the locus

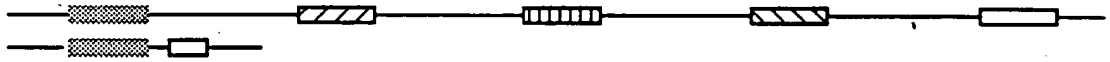
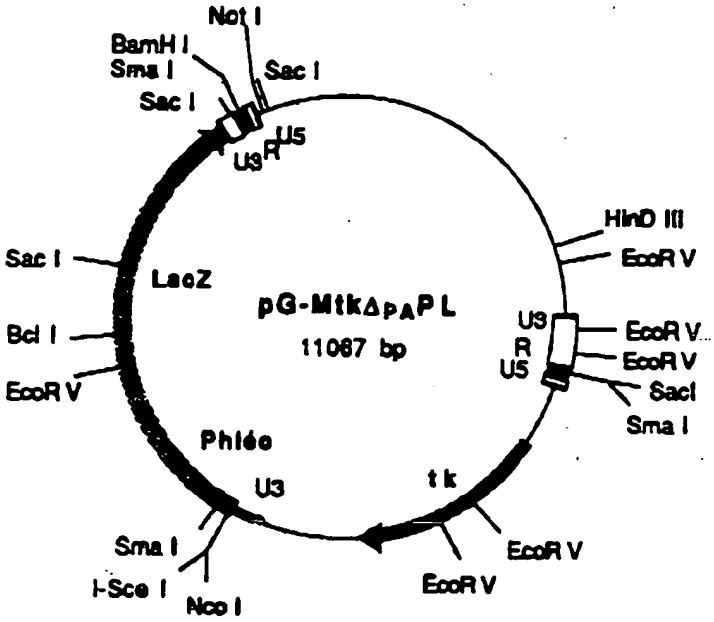


Figure 33


Not I
BamH I
Sma I
Sac I
Sac I
U3
R
U5
HinD III
EcoR V
Sac I
LacZ
pG-MtkΔpAP L
11087 bp
Bcl I
EcoR V
U3
R
U5
EcoR V
EcoR V
SacI
Sma I
Phléo
U3
tk
Sma I
I-Sce I
Nco I
EcoR V
EcoR V